

**WO 00/78970 A1**



(74) Mandataire: BOUVET, Philippe; Aventis Pharma S.A.,  
Direction Brevets, 20, avenue Raymond Aron, F-92165  
Antony Cedex (FR).

(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,  
MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM,  
GA, GN, GW, ML, MR, NE, SN, TD, TG).

(81) États désignés (*national*): AE, AG, AL, AM, AT, AU, AZ,  
BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK,  
DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID,  
IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT,  
LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ,  
PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT,  
TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

**Publiée:**

- Avec rapport de recherche internationale.
- Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues.

(84) États désignés (*régional*): brevet ARIPO (GH, GM, KE,  
LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), brevet eurasien  
(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

---

(57) Abrégé: La présente invention concerne des acides nucléiques correspondant aux différents exons et introns du gène ABC1, dont il est présentement démontré qu'il est un gène causal de pathologies liées à un dysfonctionnement du métabolisme du cholestérol induisant des maladies comme l'athérosclérose, plus particulièrement des perturbations du transport inverse du cholestérol, et plus particulièrement des déficiences familiales en HDL (FHD), comme la maladie de Tangier.

## ACIDES NUCLEIQUES ET PROTEINES CORRESPONDANT AU GENE ABC1 HUMAIN

La présente invention concerne des acides nucléiques correspondant aux différents exons et introns du gène ABC1, dont il est présentement démontré qu'il est un gène causal de pathologies liées à un dysfonctionnement du métabolisme du cholestérol induisant des maladies comme l'athérosclérose, plus particulièrement des perturbations du transport inverse du cholestérol, et plus particulièrement des déficiences familiales en HDL (FHD), comme la maladie de Tangier. L'invention concerne également des moyens de détection de polymorphismes en général, et de mutations en particulier, dans le gène ABC1 ou dans la protéine correspondante produite par la forme allélique du gène ABC1. L'invention fournit également des compositions pharmaceutiques comprenant un acide nucléique contenant la région codante du gène ABC1 et des compositions pharmaceutiques contenant la protéine ABC1 destinées au traitement de maladies liées à un déficit dans le transport inverse du cholestérol, telle que la maladie de Tangier. L'invention fournit également des méthodes de criblages de petites molécules agissant sur la protéine ABC1 qui peuvent par elle même constituer des produit agissant sur le transport inverse du cholestérol et en tant que telles, peuvent permettre de lutter efficacement contre l'athérosclérose du point de vue thérapeutique.

Les lipoprotéines de haute densité (HDL) sont l'une des quatre classes majeures de lipoprotéines qui circulent dans le plasma sanguin.

Ces lipoprotéines sont impliquées dans différentes voies métaboliques telles que le transport lipidique, la formation des acides biliaires, la stéroïdogénèse, la prolifération cellulaire et en outre interfèrent avec les systèmes de protéinase plasmatique.

Les HDL sont de parfaits accepteurs de cholestérol libre et, en combinaison avec les protéines de transfert d'ester de cholestérol (CETP), la lipoprotéine lipase (LPL), la lipase hépatique (HL) et la lécithine : cholestérol acyltransférase (LCAT), jouent un rôle majeur dans le transport inverse du cholestérol, c'est à dire le transport du cholestérol en excès dans les cellules périphériques vers le foie pour son élimination de l'organisme sous forme d'acide biliaire.

Il a été démontré que les HDL jouent un rôle central dans le transport du cholestérol des tissus périphériques vers le foie.

Diverses maladies liées à une déficience en HDL ont été décrites, comprenant la maladie de Tangier, la déficience en HDL et la déficience en LCAT.

La déficience impliquée dans la maladie de Tangier est reliée à un déficit cellulaire dans la translocation du cholestérol cellulaire qui entraînent une dégradation des HDLs. Néanmoins, pour la maladie de Tangier, la nature exacte du déficit n'a pas encore été précisément définie.

Dans la maladie de Tangier, ce déficit cellulaire conduit à une perturbation du métabolisme lipoprotéique. Les particules HDL n'incorporant pas de cholestérol à partir des cellules périphériques et ne pouvant pas être métabolisées correctement, sont éliminées rapidement de l'organisme. La concentration plasmatique en HDL de ces patients est donc extrêmement réduite et les HDL n'assurent plus le retour du cholestérol vers le foie. Ce cholestérol s'accumule dans ces cellules périphériques et provoquent des manifestations cliniques caractéristiques telles que la formation d'amygdales orangées. De plus, d'autres perturbations lipoprotéiques comme une surproduction de triglycérides ainsi qu'une synthèse et un catabolisme intracellulaire accrus des phospholipides sont observées.

La maladie de Tangier, dont les symptômes ont été décrits ci-dessus, est classée parmi les affections familiales liées au métabolisme des HDL qui sont les plus couramment détectées chez les patients affectés de maladies coronariennes.

De nombreuses études ont montré qu'un niveau réduit de cholestérol HDL est un excellent facteur de risque permettant de dépister une affection coronarienne.

Dans ce contexte, des syndromes liés aux déficiences en HDL ont présenté un intérêt accru durant la décennie passée du fait qu'elles permettent d'accroître la compréhension du rôle des HDL dans l'athérogénèse.

Plusieurs mutations dans le gène apo A-I ont été caractérisées. Ces mutations sont rares et peuvent conduire à une absence de production d'apo A-I.

Des mutations dans les gènes codant pour la lipoprotéine lipase (LPL) ou son activateur apoC-II sont associées à des hypertriglycérémies sévères et des niveaux de HDL-c fortement réduits.

Des mutations dans le gène codant pour l'enzyme lécithine: cholestérol, acyltransférase (LCAT) sont également associées à une déficience sévère en HDL.

De plus, des dysfonctionnements dans le transport inverse du cholestérol pourraient être induits par des déficits physiologiques affectant une ou plusieurs des étapes de transport du cholestérol stocké, des vésicules intracellulaires vers la surface membranaire au niveau de laquelle celui-ci est pris en charge par les HDL.

Il existe donc un besoin croissant dans l'état de la technique d'identifier des gènes impliqués dans l'une quelconque des étapes du métabolisme du cholestérol et/ou des lipoprotéines, et en particulier de gènes associés à des dysfonctionnements du transport inverse du cholestérol des cellules périphériques vers le foie.

Récemment, une étude de la ségrégation de différentes formes alléliques de 343 marqueurs microsatellites répartis sur l'ensemble du génome et distants entre eux en moyenne de 10,3 cM a été réalisée.

L'étude de liaison (linkage) a porté sur une famille bien caractérisée sur onze générations, dont de nombreux membres sont affectés par la maladie de Tangier; la famille comportant cinq lignées de consanguinité.

Cette étude a permis d'identifier une région localisée dans le locus 9q31 du chromosome 9 humain statistiquement associé à l'affection (Rust S. et al., Nature Genetics, vol. 20, Septembre 1998, pages 96-98).

Toutefois, l'étude de RUST et al. définit seulement une large région du génome dont des altérations sont susceptibles d'être associées à la maladie de Tangier. Il est simplement précisé que la région 9q31-34 concernée contient des ESTs mais aucun gène connu.

Il a désormais été montré selon l'invention qu'une région d'environ 1cM située dans le locus 9q31 chez l'homme était associée, de manière générale, à des déficiences familiales en HDL.

De manière plus précise, il a été montré qu'un gène codant pour une protéine de la famille des transporteurs ABC, localisé précisément dans la région de 1 cM du locus 9q31, était impliqué dans des pathologies liées à un déficit dans le transport inverse du cholestérol.

Plus particulièrement, il a été montré selon l'invention que le gène codant pour le transporteur ABC-1 était muté chez des patients affectés dans le transport inverse du cholestérol, et tout particulièrement chez des patients atteints de la maladie de Tangier.

5 Les protéines transporteurs ABC ("ATP-binding cassette") constituent une famille de protéines qui sont extrêmement conservées au cours de l'évolution, de la bactérie à l'homme.

Les protéines transporteurs ABC sont impliqués dans le transport membranaire de divers substrats, par exemple des ions, des acides aminés,  
10 des peptides, des sucres, des vitamines ou encore des hormones stéroïdes.

La caractérisation de la séquence complète en acides aminés de certains transporteurs ABC a permis de déterminer que ces protéines avaient une structure générale commune, notamment deux repliements de liaison aux nucléotides (Nucleotide Binding Fold ou NBF) avec des motifs de  
15 type Walker A et B ainsi que deux domaines transmembranaires, chacun des domaines transmembranaires étant constitué de six hélices. La spécificité des transporteurs ABC pour les différentes molécules transportées apparaît être déterminée par la structure des domaines transmembranaires, alors que l'énergie nécessaire à l'activité de transport  
20 est fournie par la dégradation de l'ATP au niveau du repliement NBF.

Plusieurs des protéines transporteurs ABC qui ont été identifiées chez l'homme, ont été associées à diverses maladies.

Par exemple, la mucoviscidose est provoquée par des mutations dans le gène CFTR (cystic fibrosis transmembrane conductance regulator).

25 Par ailleurs, certains phénotypes de résistance multiple aux médicaments dans les cellules tumorales ont été associés à des mutations dans le gène codant la protéine MDR (multi-drug resistance), qui a également une structure de transporteur ABC.

D'autres transporteurs ABC ont été associés à des affections  
30 neuronales et tumorales (brevet US n°5,858,719) ou encore potentiellement impliqués dans des maladies provoquées par une altération de l'homéostasie des métaux, telle que la protéine ABC-3.

De même, un autre ABC transporteur, désigné PFIC2, semble impliquer dans une forme de cholestasie intra hépatique familiale

progressive, cette protéine étant potentiellement responsable, chez l'homme, de l'exportation des sels biliaires.

En 1994, un ADNc codant pour un nouveau transporteur ABC de souris a été identifié et désigné ABC1 (Luciani et al., 1994). Cette protéine  
5 est caractéristique des transporteurs ABC en ce qu'elle comporte une structure symétrique comprenant deux domaines transmembranaires liés à un segment hautement hydrophobe et à deux motifs NBF.

Chez l'homme, un ADNc partiel comprenant la totalité de la phase de lecture ouverte du transporteur ABC1 humain a été identifié (Langmann et  
10 al., 1999).

Il a également été montré que le gène codant pour la protéine ABC1 humaine est exprimé dans divers tissus, et plus particulièrement à des niveaux élevés dans le placenta, le foie, le poumon, les glandes surrénales ainsi que les tissus fœtaux.

15 Ces auteurs ont également montré que l'expression du gène codant pour la protéine ABC1 humaine était induite pendant la différenciation des monocytes en macrophages *in vitro*. De plus, l'expression du gène codant pour la protéine ABC1 est augmentée lorsque les macrophages humains sont incubés en présence de lipoprotéines de faible densité acétylées  
20 (AcLDLs).

Toutefois, le rôle exact de la protéine ABC1 humaine dans le système de transport des lipides est totalement inconnu. Il est simplement supposé que la protéine ABC1 possède une activité de translocase des phospholipides.

25 Il a désormais été montré selon l'invention que des patients atteints de la maladie de Tangier comportaient un gène ABC1 muté. Plusieurs mutations distribuées dans différents exons du gène ABC1 ont été identifiées dans le génome de différents malades, en particulier de malades affectés d'une forme sévère de la maladie associée à des désordres  
30 coronariens. Par ailleurs, divers polymorphismes ont été trouvés à la fois dans les exons et dans les introns du gène ABC1 chez des patients atteints de formes plus légères de la maladie, indiquant que ces patients portent des allèles particuliers du gène, distinct du ou des allèles "sauvages". De tels allèles, en partie caractérisables par ces polymorphismes, sont par ailleurs  
35 susceptibles de contenir des substitutions, additions ou délétions de

nucléotides dans des régions non codantes localisées respectivement du côté 5' du premier exon ou encore du côté 3' du dernier exon du gène, en particulier dans des régions régulatrices, par exemple dans des séquences promotrices ou encore dans des séquences activatrices (en anglais "enhancer"), de nature à induire des défauts-augmentation ou diminution-  
5 dans la synthèse du polypeptide ABC1.

Il a ainsi été identifié une première mutation particulière chez un patient atteint de la maladie de Tangier, dans le gène ABC-1, qui est localisée dans l'exon 13, et qui consiste en une substitution d'un nucléotide  
10 provoquant l'introduction d'un codon d'arrêt de traduction précoce dans la phase ouverte de lecture, conduisant à la synthèse d'un polypeptide tronqué comprenant environ d'un quart de la séquence d'acides aminés du polypeptide synthétisé chez les patients non affectés par la maladie de Tangier.

15 Une seconde mutation particulière dans le gène ABC1 a été trouvée, qui consiste en une insertion d'un fragment de 100 nucléotides dans l'exon 12, conduisant à la synthèse d'un polypeptide anormal en ce qu'il contient une délétion de 6 résidus et une insertion de 38 amino-acides, ce en position 468 de la séquence de la protéine.

20 Il a en outre été confirmé selon l'invention que le gène ABC1 était régulé positivement par les lipoprotéines de faible densité acétylées (AcLDLs).

## DEFINITIONS GENERALES

25

Le terme "isolé" au sens de la présente invention désigne un matériel biologique (acide nucléique ou protéine) qui a été soustrait à son environnement originel (l'environnement dans lequel il est localisé naturellement).

30 Par exemple un polynucléotide présent à l'état naturel dans une plante ou un animal n'est pas isolé. Le même polynucléotide séparé des acides nucléiques adjacents au sein desquels il est naturellement inséré dans le génome de la plante ou l'animal est considéré comme "isolé".

Un tel polynucléotide peut être inclus dans un vecteur et/ou un tel  
35 polynucléotide peut être inclus dans une composition et demeurer



néanmoins à l'état isolé du fait que le vecteur ou la composition ne constitue pas son environnement naturel.

Le terme "purifié" ne nécessite pas que le matériel soit présent sous une forme de pureté absolue, exclusive de la présence d'autres composés. Il s'agit plutôt d'une définition relative.

Un polynucléotide est à l'état "purifié" après purification du matériel de départ ou du matériel naturel d'au moins un ordre de grandeur, de préférence 2 ou 3 et préférentiellement 4 ou 5 ordres de grandeur.

Aux fins de la présente description, l'expression "séquence nucléotidique" peut être employée pour désigner indifféremment un polynucléotide ou un acide nucléique. L'expression "séquence nucléotidique" englobe le matériel de génétique lui-même et n'est donc pas restreinte à l'information concernant sa séquence.

Les termes "acide nucléique", "polynucléotide", "oligonucléotide" ou encore "séquence nucléotidique" englobent des séquences d'ARN, d'ADN, d'ADNc ou encore des séquences hybrides ARN/ADN de plus d'un nucléotide, indifféremment sous la forme simple chaîne ou sous la forme de duplex.

Le terme "nucléotide" désigne à la fois les nucléotides naturels (A, T, G, C) ainsi que des nucléotides modifiés qui comprennent au moins une modification telle que (1) un analogue d'une purine, (2) un analogue d'une pyrimidine, ou (3) un sucre analogue, des exemples de tels nucléotides modifiés étant décrits par exemple dans la demande PCT N°WO 95/04 064.

Aux fins de la présente invention, un premier polynucléotide est considéré comme étant "complémentaire" d'un second polynucléotide lorsque chaque base du premier nucléotide est appariée à la base complémentaire du second polynucléotide dont l'orientation est inversée. Les bases complémentaires sont A et T (ou A et U), ou C et G.

Par "variant" d'un acide nucléique selon l'invention, on entendra un acide nucléique qui diffère d'une ou plusieurs bases par rapport au polynucléotide de référence. Un acide nucléique variant peut être d'origine naturel, tel qu'un variant allélique retrouvé naturellement, ou peut être aussi un variant non naturel obtenu par exemple par des techniques de mutagénèse.

En général, les différences entre l'acide nucléique de référence et l'acide nucléique variant sont réduites de telle sorte que les séquences nucléotidiques de l'acide nucléique de référence et de l'acide nucléique variant sont très proches et, dans de nombreuses régions, identiques. Les modifications de nucléotides présentes dans un acide nucléique variant peuvent être silencieuses, ce qui signifie qu'elles n'altèrent pas les séquences d'acides aminés codées par ledit acide nucléique variant.

Cependant, les changements de nucléotides dans un acide nucléique variant peuvent aussi résulter dans des substitutions, additions, délétions dans le polypeptide codé par l'acide nucléique variant par rapport aux peptides codés par l'acide nucléique de référence. En outre, des modifications de nucléotides dans les régions codantes peuvent produire des substitutions, conservatives ou non conservatives dans la séquence d'acides aminés.

De préférence, les acides nucléiques variants selon l'invention codent pour des polypeptides qui conservent sensiblement la même fonction ou activité biologique que le polypeptide de l'acide nucléique de référence ou encore la capacité à être reconnus par des anticorps dirigés contre les polypeptides codés par l'acide nucléique initial.

Certains acides nucléiques variants coderont ainsi pour des formes mutées des polypeptides dont l'étude systématique permettra de déduire des relations structure activité des protéines en question. La connaissance de ces variants par rapport à la maladie étudiée est fondamentale puisqu'elle permet de comprendre la cause moléculaire de la pathologie.

On entendra par "fragment" un acide nucléique de référence selon l'invention, une séquence nucléotidique de longueur réduite par rapport à l'acide nucléique de référence et comprenant, sur la partie commune, une séquence en nucléotides identique à l'acide nucléique de référence.

Un tel "fragment" d'acide nucléique selon l'invention peut être le cas échéant, compris dans un polynucléotide plus grand duquel il est constitutif.

De tels fragments comprennent, ou alternativement consistent en, des oligonucléotides de longueur allant de 8, 10, 12, 15, 18, 20 à 25, 30, 40, 50, 70, 80, 100, 200, 500, 1000 ou 1500 nucléotides consécutifs d'un acide nucléique selon l'invention.

Par "variant" d'un polypeptide selon l'invention, on entendra principalement un polypeptide dont la séquence d'acides aminés contient une ou plusieurs substitutions, additions ou délétions d'au moins un résidu d'acide aminé, par rapport à la séquence d'acides aminés du polypeptide de référence, étant entendu que les substitutions d'acides aminés peuvent être indifféremment conservatives ou non conservatives.

Par "fragment" d'un polypeptide selon l'invention, on entendra un polypeptide dont la séquence d'acides aminés est plus courte que celle du polypeptide de référence et qui comprend sur toute la partie commune avec ces polypeptides de référence, une séquence en acides aminés identique.

De tels fragments peuvent, le cas échéant, être compris au sein d'un polypeptide plus grand duquel ils font partie.

De tels fragments d'un polypeptide selon l'invention peuvent avoir une longueur de 10, 15, 20, 30 à 40, 50, 100, 200 ou 300 acides aminés.

Le "pourcentage d'identité" entre deux séquences de nucléotides ou d'acides aminés, au sens de la présente invention, peut être déterminé en comparant deux séquences alignées de manière optimale, à travers une fenêtre de comparaison.

La partie de la séquence nucléotidique ou polypeptidique dans la fenêtre de comparaison peut ainsi comprendre des additions ou des délétions (par exemple des "gaps") par rapport à la séquence de référence (qui ne comprend pas ces additions ou ces délétions) de manière à obtenir un alignement optimal des deux séquences.

Le pourcentage est calculé en déterminant le nombre de positions auquel une base nucléique ou un résidu d'acide aminé identique est observé pour les deux séquences (nucléique ou peptidique) comparées, puis en divisant le nombre de positions auquel il y a identité entre les deux bases ou résidus d'acides aminés par le nombre total de positions dans la fenêtre de comparaison, puis en multipliant le résultat par 100 afin d'obtenir le pourcentage d'identité de séquence.

L'alignement optimal des séquences pour la comparaison peut être réalisé de manière informatique à l'aide d'algorithmes connus contenus dans le package de la Société WISCONSIN GENETICS SOFTWARE PACKAGE, GENETICS COMPUTER GROUP (GCG), 575 Science Doctor, Madison, WISCONSIN.

À titre d'illustration, le pourcentage d'identité de séquence pourra être effectué à l'aide du logiciel BLAST (versions BLAST 1.4.9 de mars 1996, BLAST 2.0.4 de février 1998 et BLAST 2.0.6 de septembre 1998), en utilisant exclusivement les paramètres par défaut (S. F Altschul et al, J. Mol. Biol. 1990 215 : 403-410, S. F Altschul et al, Nucleic Acids Res. 1997 25 : 3389-3402). Blast recherche des séquences similaires/homologues à une séquence " requête " de référence, à l'aide de l'algorithme d'Altschul et al. La séquence requête et les bases de données utilisées peuvent être peptidiques ou nucléiques, toute combinaison étant possible.

10

Par " conditions d'hybridation de forte stringence " au sens de la présente invention, on entendra les conditions suivantes :

1- Compétition des membranes et PRE HYBRIDATION :

15

- Mélanger : 40µl ADN sperme de saumon (10mg/ml)  
+ 40 µl ADN placentaire humain (10mg/ml)

- Dénaturer 5 mn à 96°C, puis plonger dans la glace le mélange.

20

- Oter le SSC 2X et verser 4 ml de mix formamide dans le tube d'hybridation contenant les membranes.

- Ajouter le mélange des deux ADNs dénaturés.

25

- Incubation à 42°C pendant 5 à 6 heures, avec rotation.

2- Compétition de la sonde marquée :

30 - Ajouter à la sonde marquée et purifiée 10 à 50 µl ADN Cot I, selon la quantité de repeats.

- Dénaturer 7 à 10 mn à 95°C.

35 - Incuber à 65°C pendant 2 à 5 heures.

### 3- HYBRIDATION :

- Oter le mix de pré hybridation.

5

- Mélanger 40 µl ADN sperme de saumon + 40 µl ADN placentaire humain ;  
dénaturer 5 mn à 96°C, puis plonger dans la glace.

10 - Ajouter dans le tube d'hybridation 4 ml de mix formamide, le mélange des  
deux ADN et la sonde marquée/ADN Cot I dénaturée.

- Incuber 15 à 20 heures à 42°C, avec rotation.

### 4- Lavages :

15

- Un lavage à température ambiante dans du SSC 2X, pour rincer.

- 2 fois 5 minutes à température ambiante SSC 2X et SDS 0,1% à 65°C.

- 2 fois 15 minutes à 65°C SSC 1X et SDS 0,1% à 65°C.

20 Envelopper les membranes dans du Saran et exposer.

25 Les conditions d'hybridation décrites plus haut sont adaptées à  
l'hybridation dans des conditions de forte stringence, d'une molécule d'acide  
nucléique d'une longueur variable de 20 nucléotides à plusieurs centaines  
de nucléotides.

30 Il va sans dire que les conditions d'hybridation ci-dessus décrites  
peuvent être adaptées en fonction de la longueur de l'acide nucléique dont  
l'hybridation est recherchée ou du type de marquage choisi, selon des  
techniques connues de l'homme du métier.

Les conditions convenables d'hybridation peuvent par exemple être  
adaptées selon l'enseignement contenu dans l'ouvrage de HAMES et  
HIGGINS (1985) ou encore dans l'ouvrage de F.AUSUBEL et al (1999).

**ACIDES NUCLEIQUES DU GENE ABC1****SEQUENCES GENOMIQUES**

Le gène ABC1 humain comprendrait 48 exons et 47 introns, si l'on se réfère notamment à la structure du gène ABC1 orthologue chez la souris:

- 5 Plusieurs séquences nucléotidiques génomiques partielles du gène ABC1 ont été isolées et caractérisées. Selon l'invention, ces séquences génomiques comprenant à la fois des séquences exoniques et des séquences introniques nouvelles, qui peuvent être utilisées notamment pour la réalisation de différents moyens de détection du gène ABC1 ou de ses
- 10 produits d'expression nucléotidiques dans un échantillon. Ces séquences génomiques partielles sont représentées dans le tableau 1 ci-après.

**Tableau I****Séquences génomiques partielles du gène ABC1 humain**

SEQ ID NO	Désignation
1	Intron 10(p), exon 11(p)
2	Intron 11(p), exon 12, intron 12, exon 13, intron 13, exon 14, intron 14, exon 15, intron 15, exon 16, intron 16, exon 17(p)
3	Exon 17(p), intron 17(p)
4	Intron 18(p), exon 19, intron 19(p)
5	Intron 19(p), exon 20, intron 20, exon 21, intron 21, exon 22, intron 22, exon 23, intron 23, exon 24, intron 24, exon 25, intron 25, exon 26, intron 26(p)
6	Intron 26(p), exon 27, intron 27, exon 28, intron 28, exon 29, intron 29, exon 30, intron 30(p)
7	Intron 30(p), exon 31, intron 31, exon 32, intron 32, exon 33, intron 33, exon 34, intron 34(p)
8	Intron 34(p), exon 35, intron 35(p)
9	Intron 35(p), exon 36, intron 36(p)

10	Intron 36(p), exon 37, intron 37, exon 38, intron 38(p)
11	Intron 39(p), exon 40, intron 40, exon 41, intron 41(p)
12	Intron 41(p), exon 42, intron 42(p)
13	Intron 46(p), exon 47, intron 47(p)
14	Dernier exon(p), Séquence en 3' du dernier exon

Ainsi, un premier objet de l'invention consiste en un acide nucléique comprenant au moins 245 nucléotides consécutifs d'un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 1-14, ou un acide nucléique de séquence complémentaire.

L'invention concerne aussi un acide nucléique ayant au moins 80%, avantageusement 90%, de préférence 95% et de manière tout à fait préférée 98% d'identité en nucléotides avec un acide nucléique comprenant au moins 245 nucléotides consécutifs d'un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 1-14, ou un acide nucléique de séquence complémentaire.

Trente deux exons du gène ABC1 ont été caractérisés, au moins partiellement, par leur séquence nucléotidique, comme indiqué dans le Tableau II ci-dessous.

Tableau II

Exon N°	SEQ ID NO	Localisé dans SEQ ID NO	Position du nucléotide en 5'	Position du nucléotide en 3'
11 (Ext 5')	15	1	3003	3153
12	16	2	398	603
13	17	2	1124	1300
14	18	2	3087	3309
15	19	2	5055	5276
16	20	2	6337	6541

17 (Ext. 5')	21	2	7646	7660
17 (Ext 3')	22	3	1	105
19	23	4	904	1035
20	24	5	284	426
21	25	5	630	767
22	26	5	1470	1690
23	27	5	2949	3021
24	28	5	4008	4210
25	29	5	5878	5926
26	30	5	6122	6235
27	31	6	561	709
28	32	6	2359	2483
29	33	6	3714	3812
30	34	6	6848	7036
31	35	7	1183	1277
32	36	7	2587	2619
33	37	7	3744	3849
34	38	7	5323	5397
35	39	8	236	405
36	40	9	989	1166
37	41	10	545	660
38	42	10	772	916
40	43	11	435	564
41	44	11	829	949
42	45	12	589	651
47	46	13	377	620
Dernier Exon (Ext 3')	47	14	1	1237

Ainsi, l'invention est également relative à un acide nucléique  
comprenant un polynucléotide choisi dans le groupe constitué des  
séquences nucléotidiques SEQ ID NO 15-47 , ou un acide nucléique de  
5 séquence complémentaire.



Par ailleurs, trente-cinq introns du gène ABC1 ont été isolés et caractérisés, au moins partiellement. Les séquences nucléotidiques des introns du gène ABC1, ainsi que leurs fragments et leurs variants peuvent aussi être utilisés comme sondes ou amorces nucléotidiques pour détecter la présence d'au moins une copie du gène ABC1 dans un échantillon, ou encore pour amplifier une séquence cible déterminée au sein du gène ABC1.

Les références aux séquences introniques du gène ABC1 sont indiquées dans le Tableau III ci-dessous.

Tableau III

Intron N°	SEQ ID NO	Localisé dans SEQ ID NO	Position du nucléotide en 5'	Position du nucléotide en 3'
10 (Ext 3')	48	1	1	3002
11 (Ext 3')	49	2	1	397
12	50	2	604	1123
13	51	2	1301	3086
14	52	2	3310	5054
15	53	2	5277	6336
16	54	2	6542	7645
17(Ext 3')	55	3	106	1285
18 (Ext 3')	56	4	1	903
19 (Ext 5')	57	4	1036	1521
19 (Ext 3')	58	5	1	283
20	59	5	427	629
21	60	5	768	1469
22	61	5	1691	2948
23	62	5	3022	4007
24	63	5	4211	5877
25	64	5	5927	6121
26 (Ext 5')	65	5	6236	6519
26 (Ext 3')	66	6	1	560

27	67	6	710	2358
28	68	6	2484	3713
29	69	6	3813	6847
30 (Ext 5')	70	6	7037	7378
30 (Ext 3')	71	7	1	1182
31	72	7	1278	2586
32	73	7	2620	3743
33	74	7	3850	5322
34 (Ext 5')	75	7	5398	5689
34 (Ext 3')	76	8	1	235
35 (Ext 5')	77	8	406	645
35 (Ext 3')	78	9	1	988
36 (Ext 5')	79	9	1167	1664
36 (Ext 3')	80	10	1	544
37	81	10	661	771
38 (Ext 5')	82	10	917	1279
39 (Ext 3')	83	11	1	434
40	84	11	565	828
41 (Ext 5')	85	11	950	1124
41 (Ext 3')	86	12	1	588
42 (Ext 5')	87	12	652	729
46 (Ext 3')	88	13	1	376
47 (Ext 5')	89	13	621	731
Séquence distale en 3' du dernier Exon	90	14	1238	3501

L'invention a également trait à un acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 48-89, ou un acide nucléique de séquence complémentaire.

L'invention a en outre pour objet un acide nucléique ayant au moins 80% d'identité en nucléotides avec un polynucléotide choisi dans le groupe

constitué des séquences nucléotidiques SEQ ID NO 48-89, ou un acide nucléique de séquence complémentaire.

L'invention est également relative à un acide nucléique hybridant, dans des conditions de forte stringence, avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 48-89, ou un acide nucléique de séquence complémentaire.

En outre, une séquence nucléotidique génomique potentiellement régulatrice localisée en aval de l'extrémité 3' du dernier exon du gène ABC1 a été isolée. Il s'agit du polynucléotide de séquence SEQ ID NO 90. La caractérisation de polymorphismes dans cette séquence potentiellement régulatrice (présence éventuelle de séquences régulatrices de type activatrice ou " enhancer "), en particulier chez des patients affectés de formes légères de déficit dans le transport inverse du cholestérol, en particulier de formes légères de la maladie de Tangier, serait de nature à permettre la réalisation de moyens de détection appropriés, sondes ou amorces, spécifiques de certains de ces polymorphismes susceptibles d'induire des défauts dans la régulation de l'expression du gène ABC1.

Afin d'identifier les fragments polynucléotidiques biologiquement actifs de la séquence SEQ ID NO 90, l'homme du métier pourra avantageusement se référer à l'ouvrage de Sambrook et al. (1989) qui décrit l'utilisation d'un vecteur recombinant portant un gène marqueur (par exemple la  $\beta$  galactosidase, la chloramphenicol acétyl transférase, etc.) dont l'expression peut être détectée lorsque ce gène marqueur est placé sous le contrôle d'un promoteur adapté et d'un fragment biologiquement actif du polynucléotide de séquence SEQ ID NO 90. De tels fragments biologiquement actifs de la séquence SEQ ID NO 90 peuvent être notamment clonés dans des vecteurs de sélection de séquences régulatrices appropriés, tels que l'un des vecteurs pSEAP-Basic, pSEAP-Enhancer, p $\beta$ gal-Basic, p $\beta$ gal-Enhancer, ou encore pEGFP-1, commercialisés par la Société Clontech.

L'invention a en outre pour objet un acide nucléique ayant au moins 80% d'identité en nucléotides avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 90, ou un acide nucléique de séquence complémentaire.

L'invention est également relative à un acide nucléique hybridant, dans des conditions de forte stringence, avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 90, ou un acide nucléique de séquence complémentaire.

5

### **ADNc COMPLET**

Comme déjà indiqué précédemment, une séquence partielle de l'ADNc correspondant à l'expression du gène ABC1 a été identifiée par  
10 Langman et al. (1999). Cette séquence partielle de l'ADNc d'ABC1 comprend 6880 nucléotides et contient la totalité de la phase ouverte de lecture correspondant à la protéine ABC1 produite chez les sujets non affectés de troubles liés au transport inverse du cholestérol. La séquence d'ADNc décrite par Langmann et al. (1999) contient en outre une partie de  
15 la région 5'-UTR (nucléotides 1 à 120) et une partie de la région 3'-UTR (nucléotides 6727 à 6880).

Il a désormais été isolé et caractérisé selon l'invention la totalité de l'ADNc complet correspondant au gène ABC1, qui comprend une région 3'-UTR nouvelle, qui constitue une région nucléique importante, notamment du  
20 point de vue de la stabilité des ARN messagers dans la cellule.

Les analyses d'expression du transcrit de séquence SEQ ID N°91 ont été réalisées par RT-PCR, comme décrit dans l'Exemple 1. Ces analyses effectuées à partir d'ARN polyA+ de différents tissus ont permis de montrer que le gène ABC1 était exprimé dans le cerveau fœtal, le cerveau, le cœur,  
25 le placenta et l'utérus.

En conséquence, l'invention concerne aussi un acide nucléique comprenant un polynucléotide de séquence SEQ ID NO 91 de l'ADNc du gène ABC1 humain, ou un acide nucléique de séquence complémentaire.

L'ADNc du gène ABC1 humain de séquence SEQ ID NO 91  
30 comprend une phase de lecture ouverte allant du nucléotide en position 121 (base A du codon d'initiation de la traduction ATG) au nucléotide en position 6723 de la séquence SEQ ID NO 91. Un signal de polyadénylation (de séquence ATTAAG) est présent, débutant au nucléotide en position 9454 de la séquence SEQ ID NO 91.

L'ADNc de séquence SEQ ID NO 91 code pour le polypeptide ABC1 d'une longueur de 2201 acides aminés, et ayant la séquence d'acides aminés SEQ ID NO 139.

- L'invention a également trait à un acide nucléique comprenant au
5. moins huit nucléotides consécutifs d'un polynucléotide de séquence SEQ ID NO 92, un fragment biologiquement actif de celui-ci ou un acide nucléique de séquence complémentaire.

- L'invention a aussi pour objet un acide nucléique ayant au moins 80% d'identité en nucléotides avec un polynucléotide de séquence SEQ ID NO
- 10 92, un fragment biologiquement actif de celui-ci ou un acide nucléique de séquence complémentaire.

- Un autre objet de l'invention consiste en un acide nucléique hybridant, dans des conditions de forte stringence, avec un polynucléotide de séquence SEQ ID NO 92, un fragment biologiquement actif de celui-ci ou un
- 15 acide nucléique de séquence complémentaire.

## **POLYMORPHISMES AU SEIN DU GENE ABC1**

### **MUTATIONS**

- Selon l'invention, plusieurs mutations ont été identifiées dans la
- 20 séquence du gène ABC1, ces mutations conduisant à des altérations structurales majeures du polypeptide ABC1 codé par les séquences mutées. Ces mutations ont été retrouvées particulièrement dans des patients atteints de formes sévères de la maladie de Tangier, associées à de graves désordres coronariens. Deux mutations particulièrement délétères sont
- 25 décrites ci-après.

#### **1. Mutation dans l'exon 12**

- Cette mutation consiste à la fois en une délétion d'un segment de 14 nucléotides (" TGAGAGGAAGTTCT ") localisé du nucléotide en position 472
- 30 au nucléotide en position 485 de l'ADN génomique normal de séquence SEQ ID NO 2 et en une insertion d'une séquence de type Alu de 110 nucléotides au sein de la séquence de l'exon12 du gène ABC1, en amont du nucléotide en position 486 de l'ADN génomique normal de séquence SEQ ID NO 2.

L'exon 12 portant cette mutation de délétion/insertion a la séquence nucléotidique SEQ ID NO 93.

L'ADNc muté correspondant a la séquence nucléotidique SEQ ID NO 94, code pour un polypeptide ABC1 muté d'une longueur de 2233 acides aminés, de séquence SEQ ID NO 140, dont la structure est fortement altérée par rapport au polypeptide ABC1 normal de séquence SEQ ID NO 139.

Les séquences nucléotidiques SEQ ID NO 93 et 94 ainsi que la séquence polypeptidique SEQ ID NO 140 font également partie de l'invention.

10

### **2 Mutation dans l'exon 13**

Cette mutation consiste en une délétion du nucléotide (G) en position 1232 de la séquence génomique SEQ ID NO 2, qui est localisé dans l'exon 13 (nucléotide G en position 106 de la séquence de l'exon 13 SEQ ID NO 17). Cette délétion ponctuelle d'une base introduit un codon stop dans la phase de lecture normale dans l'ARNm du gène ABC1.

La séquence de l'exon 13 du gène ABC1 portant cette mutation est le polynucléotide de séquence SEQ ID NO 95.

L'ADNc correspondant à cette mutation dans l'exon 13 du gène ABC1 est représenté par la séquence nucléotidique SEQ ID NO 96.

La protéine mutée codée par le gène ABC1 muté ayant une longueur de 574 acides aminés, c'est-à-dire environ le quart de la longueur en acides aminés de la protéine normale. Le polypeptide tronqué a la séquence d'acides aminés SEQ ID NO 141.

Les caractéristiques structurales permettant de différencier les séquences normales des séquences mutées d'ABC1 (génomiques, ARN messagers, ADNc) peuvent être mises à profit afin de réaliser des moyens de détections des séquences mutées d'ABC1 dans un échantillon, en particulier des sondes hybridant spécifiquement avec les séquences mutées d'ABC1 ou encore des couples d'amorces permettant d'amplifier sélectivement les régions du gène ABC1 portant les mutations décrites ci-dessus, la détection de la présence de ces mutations pouvant notamment être effectuée par discrimination de la longueur des fragments d'acide nucléique amplifiés, par hybridation des fragments amplifiés à l'aide des

30

sondes spécifiques décrites ci-dessus, ou encore par séquençage direct de ces fragments amplifiés.

Ainsi, l'invention a encore pour objet un acide nucléique ayant au moins huit nucléotides consécutifs d'un polynucléotide choisi dans le groupe  
5 constitué des séquences nucléotidiques SEQ ID NO 93-96, ou un acide nucléique de séquence complémentaire.

De préférence, un tel acide nucléique comprend :

- a) soit au moins deux nucléotides consécutifs de la séquence Alu localisée dans les séquences SEQ ID NO 93 et 94, de préférence  
10 5, 10, 15, 20, 25, 30, 35, 40, 50 ou 100 nucléotides consécutifs de la séquence Alu localisée dans les séquences SEQ ID NO 93 et 94 ;
- b) soit au moins les deux nucléotides " CT " situés de par et d'autre de la base G délétée, dans les séquences SEQ ID NO 94  
15 et 95.

Font également partie de l'invention les amorces hybridant avec une séquence nucléique localisée dans la région d'une séquence d'ABC1 (génomique, ARN messenger) portant l'une ou l'autre des deux mutations décrites ci-dessus.

20 L'invention concerne en outre un acide nucléique ayant au moins 80% d'identité en nucléotides avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 93-96, ou un acide nucléique de séquence complémentaire.

L'invention est également relative à un acide nucléique hybridant,  
25 dans des conditions de forte stringence, avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 93-96, ou un acide nucléique de séquence complémentaire.

#### **AUTRES POLYMORPHISMES**

30 D'autres polymorphismes ont été retrouvés au sein de la séquence du gène ABC1, notamment des substitutions de nucléotides localisées à la fois dans les régions codantes (exons) et dans les régions non codantes.

Concernant les polymorphismes retrouvés au sein des régions codantes, il s'agit essentiellement de substitutions d'un seul nucléotide  
35 localisé sur la troisième base des codons du cadre ouvert de lecture d'ABC1,

ces substitutions n'entraînant pas de modification quant à la nature de l'acide aminé codé, compte tenu des règles de dégénérescence génétique chez l'homme, bien connues de l'homme du métier.

Ces polymorphismes sont représentés dans la présente description sous la forme de séquences nucléotidiques d'une longueur de 41 bases, la base polymorphe étant localisée au centre du fragment polymorphe. Pour chacun des polymorphismes repérés, chaque allèle est ainsi représenté comme une séquence de 41 bases, le polymorphisme lui-même étant défini par les deux séquences nucléotidiques correspondant respectivement à chacune des formes. Les polymorphismes identifiés dans le gène ABC1 sont représentés dans le Tableau IV ci-dessous.

**Tableau IV**  
**Polymorphismes retrouvés dans le gène ABC1**

Désignation N°	Position dans la séquence SEQ ID NO 2	Allèle 1 SEQ ID NO	Allèle 2 SEQ ID NO	Base polymorphe Allèle1/Allèle 2
1	397	97	98	G/A
2	1324	99	100	T/A
3	3028	101	102	C/T
4	3234	103	104	C/A
5	3390	105	106	A/G
6	6854	107	108	G/A

La détection de ces polymorphismes au sein d'un échantillon d'ADN provenant d'un sujet peut par exemple être réalisée par une amplification spécifique de la région nucléotidique d'ABC1 contenant la base polymorphe, puis séquençage du fragment amplifié afin de déterminer la nature de l'allèle ou des allèles portés par ledit sujet.

La détection de ces polymorphismes au sein d'un échantillon d'ADN provenant d'un sujet peut aussi être réalisé à l'aide de sondes ou d'amorces nucléotidiques hybridant spécifiquement avec un allèle déterminé contenant



l'une des bases polymorphes d'un polymorphisme du gène ABC1 selon l'invention.

A titre illustratif, des amorces nucléotidiques appropriées sont par exemple des amorces dont la base à l'extrémité 3' hybride avec la base localisée immédiatement du côté 5' de la base polymorphe du fragment  
5 comprenant ledit polymorphisme. Après l'étape d'hybridation de l'amorce spécifique, une étape d'élongation avec un mélange des deux didéoxynucléotides complémentaires de la base polymorphe dudit polymorphisme, par exemple marqués différemment par fluorescence,  
10 puis une étape de détection du signal de fluorescence obtenu permet de déterminer lequel des deux didéoxynucléotides fluorescents différemment marqués a été incorporé et de déduire directement la nature de la base polymorphe présente au niveau de ce polymorphisme.

Différentes approches peuvent être utilisées pour le marquage et la  
15 détection des didéoxynucléotides. Une méthode en phase homogène basée sur le FRET ("Fluorescence resonance energy transfer") a été décrite par Chen et Kwok (1997). Selon cette méthode, des fragments amplifiés d'ADN génomique contenant des polymorphismes sont incubés avec une amorce marquée à la fluorescéine à l'extrémité 5', en présence de  
20 didéoxynucléotides triphosphate marqués et une polymérase Taq modifiée. L'amorce marquée est allongée d'une base par incorporation du didéoxynucléotide marqué spécifique de l'allèle présent sur la séquence d'ADN génomique complémentaire. A la fin de cette réaction de génotypage, les intensités de fluorescence pour les deux composés marqueurs des  
25 didéoxynucléotides marqués sont analysées directement sans séparation ni purification. L'ensemble de ces étapes peut être réalisé dans le même tube et les modifications du signal de fluorescence suivies en temps réel. Selon un autre mode de réalisation, l'amorce allongée peut être analysée par spectrométrie de masse de type MALDI-TOF. La base localisée au niveau  
30 du site polymorphique est identifiée par mesure de la masse ajoutée à l'amorce de microséquençage (Haff et Smirnov, 1997).

De telles amorces nucléotidiques peuvent par exemple être immobilisées sur un support. De plus, il est possible d'immobiliser sur un  
35 support, par exemple de manière ordonnée, de multiples amorces

spécifiques telles que décrites ci-dessus, chacune des amorces étant adaptée à la détection de l'un des polymorphismes du gène ABC1 selon l'invention.

Les polymorphismes du gène ABC1 selon l'invention sont utiles  
5 notamment comme marqueurs génétiques dans des études d'association entre la présence d'un allèle donné chez un sujet et la prédisposition de ce sujet à une pathologie donnée, particulièrement à l'une des pathologies déjà associées à la région chromosomique 9q31 préférentiellement à une pathologie liée à un dysfonctionnement du transport inverse du cholestérol.

10 Les méthodes pour l'analyse génétique de caractères (phénotypes) complexes sont de différents types (Lander and Schork, 1994). En général, les polymorphismes bialléliques selon l'invention sont utiles dans l'un quelconque des procédés décrits dans l'état de la technique destiné à démontrer une corrélation statistiquement significative entre un génotype et  
15 un phénotype. Les polymorphismes bialléliques peuvent être utilisés dans des analyses de liaison ("linkage analysis") et dans des procédés de partage d'allèles ("allele sharing"). De préférence, les polymorphismes bialléliques selon l'invention sont utilisés pour identifier des gènes associés à des caractères (phénotypes) détectables en utilisation des études d'association,  
20 une approche qui ne nécessite pas le recours à des familles affectées par le caractère, et qui permet en outre l'identification de gènes associés à des caractères complexes et sporadiques.

D'autres méthodes statistiques mettant en œuvre des polymorphismes bialléliques selon l'invention sont par exemple celles  
25 décrites par Forsell et al. (1997), Xiong et al. (1999), Horvath et al. (1998), Sham et al. (1995) ou encore Nickerson et al. (1992).

Selon un autre aspect, l'invention concerne également les séquences nucléotidiques du gène ABC1 comprenant au moins un polymorphisme biallélique tel que décrit ci-dessus.

30 Ainsi, l'invention est également relative à un acide nucléique ayant au moins huit nucléotides consécutifs d'un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 97-108 et comprenant la base polymorphe, ou un acide nucléique de séquence complémentaire.

**SONDES ET AMORCES NUCLEOTIDIQUES**

Les fragments d'acides nucléiques dérivés de l'une quelconque des séquences nucléotidiques SEQ ID N° 1-14, 15-47, 48-89, 90, 92, 94-96 et 97-108 sont utiles pour la détection de la présence d'au moins une copie  
5 d'une séquence nucléotidique du gène ABC1 ou encore d'un fragment ou d'un variant (contenant une mutation ou un polymorphisme) de cette dernière dans un échantillon.

Les sondes ou les amorces nucléotidiques selon l'invention comprennent au moins 8 nucléotides consécutifs d'un acide nucléique choisi  
10 dans le groupe constitué des séquences SEQ ID NO 1-14, 15-47, 48-89, 90, 92, 93-96 et 97-108, ou d'un acide nucléique de séquence complémentaire.

De préférence, des sondes ou amorces nucléotidiques selon l'invention auront une longueur de 10, 12, 15, 18 ou 20 à 25, 35, 40, 50, 70, 80, 100, 200, 500, 1000, 1500 nucléotides consécutifs d'un acide nucléique  
15 selon l'invention, en particulier un acide nucléique de séquence nucléotidique choisie parmi les séquences SEQ ID NO 1-14, 15-47, 48-89, 90, 92, 93-96 et 97-108, ou d'un acide nucléique de séquence complémentaire.

Alternativement, une sonde ou une amorce nucléotidique selon l'invention consistera et/ou comprendra les fragments d'une longueur de 12, 15, 18, 20, 25, 35, 40, 50, 100, 200, 500, 1000, 1500 nucléotides consécutifs d'un acide nucléique selon l'invention, plus particulièrement d'un acide nucléique choisi parmi les séquences SEQ ID N°1-14, 15-47, 48-89, 90, 92,  
25 93-96 et 97-108, ou d'un acide nucléique de séquence complémentaire.

La définition d'une sonde et d'une amorce nucléotidique selon l'invention englobe donc des oligonucléotides qui hybrident, dans les conditions d'hybridation de forte stringence définies ci-avant, avec un acide nucléique choisi parmi les séquences SEQ ID NO 1-14, 15-47, 48-89, 90,  
30 92, 93-96 et 97-108 ou avec une séquence complémentaire de ces derniers.

Des exemples d'amorces et de couples d'amorces permettant d'amplifier différentes régions du gène ABC1 sont représentés dans le Tableau V ci-dessous.

## Amorces pour l'amplification de fragments nucléiques du gène ABC1

Amorce N°	Localisée dans SEQ ID	Position dans la séquence	Séquence de l'amorce	Région d'hybridation
1	2	313-335	109	Intron 11
2	2	Comp 640-663	110	Intron 12
3	2	1005-1029	111	Intron 12
4	2	Comp 1472-1496	112	Intron 13
5	2	2930-2954	113	Intron 13
6	2	Comp 3444-3468	114	Intron 14
7	2	4988-5012	115	Intron 14
8	2	Comp 5338-5362	116	Intron 15
9	2	6240-6262	117	Intron 15
10	2	Comp 6581-6603	118	Intron 16
11	5	1369-1391	119	Intron 21
12	5	Comp 1748-1770	120	Intron 22
13	5	3868-3890	121	Intron 23
14	5	Comp 4240-4262	122	Intron 24
15	6	3587-3610	123	Intron 28
16	6	Comp 3881-3903	124	Intron 29
17	6	6753-6775	125	Intron 29
18	6	Comp 7112-7134	126	Intron 30
19	7	1060-1082	127	Intron 30
20	7	Comp 1377-1399	128	Intron 31

21	7	3574-3596	129	Intron 32
22	7	Comp 3909-3931	130	Intron 33
23	7	5161-5183	131	Intron 33
24	7	Comp 5463-5485	132	Intron 34
25	8	100-122	133	Intron 34
26	8	Comp 475-497	134	Intron 35
27	9	841-861	135	Intron 35
28	9	Comp 1249-1271	136	Intron 36
29	10	455-477	137	Intron 36
30	10	Comp 966-988	138	Intron 38

Selon un premier mode de réalisation de sondes et d'amorces préférées selon l'invention, celles-ci comprennent tout ou partie d'un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID NO 109-138, ou des acides nucléiques de séquence complémentaire.

Une amorce ou une sonde nucléotidique selon l'invention peut être préparée par toute méthode adaptée bien connue de l'homme du métier, y compris par clonage et action d'enzymes de restriction ou encore par synthèse chimique directe selon des techniques telles que la méthode au phosphodiester de Narang et al. (1979) ou de Brown et al. (1979), la méthode aux diéthylphosphoramidites de Beaucage et al. (1980) ou encore la technique sur support solide décrite dans le brevet EU N°EP 0 707 592.

Chacun des acides nucléiques selon l'invention, y compris les sondes et amorces oligonucléotidiques décrites ci-dessus, peut être marqué, si désiré, en incorporant un marqueur détectable par des moyens spectroscopiques, photochimiques, biochimiques, immunochimiques ou encore chimiques.

Par exemple, de tels marqueurs peuvent consister en des isotopes radioactifs ( $^{32}\text{P}$ ,  $^{33}\text{P}$ ,  $^3\text{H}$ ,  $^{35}\text{S}$ ), des molécules fluorescentes (5-bromodeoxyuridine, fluorescéine, acétylaminofluorène, digoxigénine) ou encore des ligands tels que la biotine.

Le marquage des sondes est fait de préférence par incorporation de molécules marquées au sein des polynucléotides par extension d'amorces, ou bien par rajout sur les extrémités 5' ou 3'.

Des exemples de marquage non radioactifs de fragments d'acides nucléiques sont décrits notamment dans le brevet français n° FR 78 109 75 ou encore dans les articles de Urdea et al. (1988) ou Sanchez-pescador et al. (1988).

De manière avantageuse, les sondes selon l'invention peuvent avoir des caractéristiques structurales de nature à permettre une amplification du signal, telles que les sondes décrites par Urdea et al. (1991) ou encore dans le brevet européen n° EP-0 225 807 (CHIRON).

Les sondes oligonucléotides selon l'invention peuvent être utilisées notamment dans des hybridations de type Southern à l'ADN génomique ou encore dans des hybridations à l'ARN messager correspondant lorsque l'expression du transcrit correspondant est recherchée dans un échantillon.

Les sondes selon l'invention peuvent aussi être utilisées pour la détection de produits d'amplification PCR ou encore pour la détection de mésappariements.

Des sondes ou amorces nucléotidiques selon l'invention peuvent être immobilisées sur un support solide. De tels supports solides sont bien connus de l'homme du métier et comprennent des surfaces des puits de plaques de microtitration, des lits de polystyrène, des lits magnétiques, des bandes de nitrocellulose, ou encore des microparticules telles que des particules de latex.

En conséquence, la présente invention concerne également un procédé de détection de la présence d'un acide nucléique tel que décrit ci-avant dans un échantillon, ladite méthode comprenant les étapes de :

- 1) mettre en contact une ou plusieurs sondes nucléotidiques selon l'invention avec l'échantillon à tester ;
- 2) détecter le complexe éventuellement formé entre la ou les sondes et l'acide nucléique présent dans l'échantillon.

Selon un mode de réalisation particulier du procédé de détection selon l'invention, la ou les sondes oligonucléotidiques sont immobilisées sur un support.

Selon un autre aspect, les sondes oligonucléotidiques comprennent un marqueur détectable.

L'invention concerne en outre un nécessaire ou kit pour la détection de la présence d'un acide nucléique selon l'invention dans un échantillon, 5 ledit nécessaire comprenant :

- a) une ou plusieurs sondes nucléotidiques telles que décrites ci-dessus ;
- b) le cas échéant, les réactifs nécessaires à la réaction d'hybridation.

10 Selon un premier aspect, le nécessaire ou kit de détection est caractérisé en ce que la ou les sondes sont immobilisées sur un support.

Selon un second aspect, le nécessaire ou kit de détection est caractérisé en ce que les sondes oligonucléotidiques comprennent un marqueur détectable.

15 Selon un mode de réalisation particulier du kit de détection décrit ci-dessus, un tel kit comprendra une pluralité de sondes oligonucléotidiques conformes à l'invention qui pourront être utilisées pour détecter des séquences cibles d'intérêt ou alternativement détecter des mutations dans les régions codantes ou les régions non codantes des acides nucléiques 20 selon l'invention, plus particulièrement des acides nucléiques de séquences SEQ ID NO 1-14, 15-47, 48-89, 90, 92, 93-96 et 97-108 ou les acides nucléiques de séquence complémentaire.

Ainsi, les sondes selon l'invention immobilisées sur un support peuvent être ordonnées en matrices telles que les " puces à ADN ". De telles 25 matrices ordonnées ont été en particulier décrites dans le brevet US N° 5,143,854, dans les demandes PCT N° WO 90/150 70 et 92/10092.

Des matrices supports sur lesquelles des sondes oligonucléotidiques ont été immobilisées à une haute densité sont par exemple décrites dans les brevets US N°5,412,087 et dans la demande PCT N°WO 95/11995.

30 Les amorces nucléotidiques selon l'invention peuvent être utilisées pour amplifier l'un quelconque des acides nucléiques selon l'invention, et plus particulièrement tout ou partie d'un acide nucléique de séquences SEQ ID NO 1-14, 15-47, 48-89, 90, 92, 93-96 et 97-108, ou encore un variant de celui-ci.

Un autre objet de l'invention concerne un procédé pour l'amplification d'un acide nucléique selon l'invention, et plus particulièrement un acide nucléique de séquences SEQ ID NO 1-14, 15-47, 48-89, 90, 92, 93-96 et 97-108 ou un fragment ou un variant de celui-ci contenu dans un échantillon, ledit procédé comprenant les étapes de :

- a) mettre en contact l'échantillon dans lequel la présence de l'acide nucléique cible est suspectée avec une paire d'amorces nucléotidiques dont la position d'hybridation est localisée respectivement du côté 5' et du côté 3' de la région de l'acide nucléique cible dont l'amplification est recherchée, en présence des réactifs nécessaires à la réaction d'amplification ; et
- b) détection des acides nucléiques amplifiés.

Pour mettre en œuvre le procédé d'amplification tel que défini ci-dessus, on aura avantageusement recours à l'une quelconque des amorces nucléotidiques décrites ci-avant.

L'invention a en outre pour objet un nécessaire ou kit pour l'amplification d'un acide nucléique selon l'invention, et plus particulièrement tout ou partie d'un acide nucléique de séquences SEQ ID NO 1-14, 15-47, 48-89, 90, 92, 93-96 et 97-108 ledit nécessaire ou kit comprenant :

- a) un couple d'amorces nucléotidiques conformes à l'invention, dont la position d'hybridation est localisée respectivement du côté 5' et du côté 3' de l'acide nucléique cible dont l'amplification est recherchée ;
- b) le cas échéant, les réactifs nécessaires à la réaction d'amplification.

Un tel nécessaire ou kit d'amplification comprendra avantageusement au moins une paire d'amorces nucléotidiques telles que décrites ci-dessus.

Selon un premier mode de réalisation préféré, des amorces selon l'invention comprennent tout ou partie d'un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID NO 109 et 110, permettant d'amplifier la région de l'exon 12 du gène ABC1 portant la première mutation (délétion/insertion) décrite ci-dessus, ou des acides nucléiques de séquence complémentaire.



Selon un second mode de réalisation préféré, des amorces selon l'invention comprennent tout ou partie d'un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID NO 111 et 112, permettant d'amplifier la région de l'exon 13 du gène ABC1 portant la seconde mutation (délétion d'une base G) décrite ci-dessus, ou des acides nucléiques de séquence complémentaire.

Selon un troisième mode de réalisation préféré, des amorces selon l'invention comprennent, de manière générale, tout ou partie d'un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID NO 109-138, ou des acides nucléiques de séquence complémentaire.

Selon un quatrième mode de réalisation préféré, l'invention a également trait à des amorces nucléotidiques comprenant au moins 15 nucléotides consécutifs d'un acide nucléique choisi dans le groupe constitué des séquences SEQ ID NO 97-108 ou un acide nucléique de séquence complémentaire, la base de l'extrémité 3' de ces amorces étant complémentaire du nucléotide localisée immédiatement du côté 5' de la base polymorphe de l'une des séquences SEQ ID NO 97-108 ou de leurs séquences complémentaires.

Selon un autre aspect, l'invention concerne aussi des amorces nucléotidiques comprenant au moins 15 nucléotides consécutifs d'un acide nucléique choisi dans le groupe constitué des séquences SEQ ID NO 97-108 ou un acide nucléique de séquence complémentaire, la base de l'extrémité 3' de ces amorces étant complémentaire d'un nucléotide situé à 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 nucléotides ou plus du côté 5' de la base polymorphe de l'une des séquences SEQ ID NO 97-108 ou de leurs séquences complémentaires. Pour construire des amorces dont le nucléotide à l'extrémité 3' est complémentaire d'un nucléotide localisé à plus de 20 nucléotides du côté 5' de la base polymorphe de l'une des séquences SEQ ID NO 97-108, l'homme du métier se référera avantageusement à la séquence génomique correspondante parmi les séquences SEQ ID NO 1-14 ou encore SEQ ID NO 15-47 et 48-90, comprenant le polymorphisme dont la nature de l'allèle est recherchée.

De telles amorces sont particulièrement utiles dans le cadre de procédés de génotypage de sujets et/ou de génotypage de populations, notamment dans le cadre d'études d'association entre des formes allèles

particulières ou des formes de groupes d'allèles particulières (haplotypes) chez des sujets et l'existence d'un phénotype (caractère) particulier chez ces sujets, par exemple la prédisposition de ces sujets à développer des maladies liées à un déficit dans le transport inverse du cholestérol, ou  
5 encore la prédisposition de ces sujets à développer une pathologie dont la région chromosomique candidate se situe sur le chromosome 9, plus précisément sur le bras 9q et plus précisément encore dans le locus 9q31.

### VECTEURS RECOMBINANTS

10 L'invention est également relative à un vecteur recombinant comprenant un acide nucléique selon l'invention.

Avantageusement, un tel vecteur recombinant comprendra un acide nucléique choisi parmi les acides nucléiques suivants :

- a) un acide nucléique de séquence SEQ ID NO 92 ou un fragment  
15 biologiquement actif de ce dernier ;
- b) un acide nucléique comprenant un polynucléotide de séquence SEQ ID NO 91, 94 ou 96 ;
- c) un acide nucléique comprenant un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 15-47 et 48-90
- 20 d) un acide nucléique ayant au moins 80% d'identité en nucléotides avec un acide nucléique choisi dans le groupe constitué des séquences SEQ ID NO 15-47 et 48-90 ou un fragment ou un variant de ce dernier ;
- d) un acide nucléique hybridant, dans des conditions d'hybridation de forte stringence, avec un acide nucléique de séquences SEQ ID NO 15-47 et  
25 48-90, ou un fragment ou un variant de ce dernier.

Par "vecteur" au sens de la présente invention on entendra une molécule d'ADN ou d'ARN circulaire ou linéaire qui est indifféremment sous forme de simple brin ou double brin.

30 Selon un premier mode de réalisation, un vecteur recombinant selon l'invention est utilisé afin d'amplifier l'acide nucléique qui y est inséré après transformation ou transfection de l'hôte cellulaire désiré.

Selon un second mode de réalisation, il s'agit de vecteurs d'expression comprenant, outre un acide nucléique conforme à l'invention,

des séquences régulatrices permettant d'en diriger la transcription et/ou la traduction.

Selon un mode de réalisation avantageux, un vecteur recombinant selon l'invention comprendra notamment les éléments suivants :

- 5 (1) des éléments de régulation de l'expression de l'acide nucléique à insérer, tels que des promoteurs et des séquences activatrices (" enhancers ") ;
- (2) la séquence codante comprise dans l'acide nucléique conforme à l'invention à insérer dans un tel vecteur, ladite séquence codante étant  
10 placée en phase avec les signaux de régulation décrits aux (1) ; et
- (3) des séquences d'initiation et d'arrêt de la transcription appropriées.

En outre, les vecteurs recombinants selon l'invention pourront inclure une ou plusieurs origines de réplication chez les hôtes cellulaires dans  
15 lesquels leur amplification ou leur expression est recherchée, des marqueurs ou des marqueurs de sélection.

A titre d'exemples, les promoteurs bactériens pourront être les promoteurs LacI, LacZ, les promoteurs de l'ARN polymérase du bactériophage T3 ou T7, les promoteurs PR, ou PL du phage lambda.

20 Les promoteurs pour cellules eucaryotes comprendront le promoteur de la thymidine kinase du virus HSV ou encore le promoteur de la métallothionéine-L de souris.

De manière générale, pour le choix d'un promoteur adapté, l'homme du métier pourra avantageusement se référer à l'ouvrage de Sambrook et al.  
25 (1989) précité ou encore aux techniques décrites par Fuller et al. (1996).

Lorsque l'expression de la séquence génomique du gène ABC1 sera désirée, on aura préférentiellement recours à des vecteurs capables d'inclure de grandes séquences d'insertion. Dans ce mode de réalisation particulier, on utilisera de préférence des vecteurs de bactériophages, tels  
30 que les vecteurs de bactériophage P1 comme le vecteur p158 ou encore le vecteur p158/neo8 décrits par Sternberg (1992, 1994).

Les vecteurs bactériens préférés selon l'invention sont par exemple les vecteurs pBR322(ATCC37017) ou encore des vecteurs tels que pAA223-3 (Pharmacia, Uppsala, Suède), et pGEM1 (Promega Biotech, Madison, WI,  
35 ETATS-UNIS).

On peut encore citer d'autres vecteurs commercialisés tels que les vecteurs pQE70, pQE60, pQE9 (Qiagen), psiX174, pBluescript SA, pNH8A, pNH16A, pNH18A, pNH46A, pWLNEO, pSV2CAT, pOG44, pXTI, pSG(Stratagene).

- 5 Il peut s'agir également de vecteurs de type *baculovirus* tel que le vecteur pVL1392/1393 (Pharmingen) utilisé pour transfecter les cellules de la lignée Sf9 (ATCC N°CRL 1711) dérivées de *Spodoptera frugiperda*.

Il peut encore s'agir de vecteurs adénoviraux tels que l'adénovirus humain de type 2 ou 5.

- 10 Un vecteur recombinant selon l'invention peut aussi être un vecteur rétroviral ou encore un vecteur adéno-associé (AAV). De tels vecteurs adéno-associés sont par exemple décrits par Flotte et al. (1992), Samulski et al. (1989), ou encore McLaughlin BA et al. (1996).

- 15 Pour permettre l'expression des polynucléotides selon l'invention, ces derniers doivent être introduits dans une cellule hôte. L'introduction des polynucléotides selon l'invention dans une cellule hôte peut être réalisée *in vitro*, selon les techniques bien connues de l'homme du métier pour transformer ou transfecter des cellules, soit en culture primaire, soit sous la  
20 forme de lignées cellulaires. On peut aussi réaliser l'introduction des polynucléotides selon l'invention *in vivo* ou *ex vivo*, pour la prévention ou le traitement de maladies liées à un déficit dans le transport inverse du cholestérol.

- Pour introduire les polynucléotides ou les vecteurs dans une cellule  
25 hôte, l'homme du métier pourra avantageusement se référer à différentes techniques, comme la technique de précipitation au phosphate de calcium (Graham et al., 1973 ; Chen et al., 1987), le DEAE Dextran (Gopal, 1985), l'électroporation (Tur-Kaspa, 1896 ; Potter et al., 1984), la microinjection directe (Harland et al., 1985), Les liposomes chargés en ADN (Nicolau et al.,  
30 1982, Fraley et al., 1979).

- Une fois que le polynucléotide a été introduit dans la cellule hôte, il peut être intégré de manière stable dans le génome de la cellule. L'intégration peut être réalisée à un endroit précis du génome, par recombinaison homologue, ou peut être intégré au hasard. Dans certains  
35 modes de réalisation, le polynucléotide peut être maintenu de manière stable

dans la cellule hôte sous la forme d'un fragment d'épisome, l'épisome comprenant des séquences permettant le maintien et la réplication de ce dernier, soit de manière indépendante, soit de manière synchronisée avec le cycle cellulaire.

5        Selon un mode de réalisation particulier, une méthode pour introduire un polynucléotide selon l'invention dans une cellule hôte, en particulier une cellule hôte provenant d'un mammifère, *in vivo*, comprend une étape au cours de laquelle on introduit une préparation comprenant un vecteur pharmaceutiquement compatible et un polynucléotide "nu" selon l'invention,  
10       placé sous le contrôle de séquences de régulation appropriées, par injection locale au niveau du tissu choisi, par exemple un tissu musculaire lisse, le polynucléotide "nu" étant absorbé par les cellules de ce tissu.

Des compositions pour l'utilisation *in vitro* et *in vivo* comprenant des polynucléotides "nus" sont par exemples décrites dans la demande PCT N°  
15       WO 95/11307 (Institut Pasteur, Inserm, Université d'Ottawa) ainsi que dans les articles de Tacson et al. (1996) et de Huygen et al. (1996).

Selon un mode de réalisation spécifique de l'invention, il est fourni une composition pour la production *in vivo* de la protéine ABC1. Cette composition comprend un polynucléotide codant pour le polypeptide ABC1  
20       placé sous le contrôle de séquences régulatrices appropriées, en solution dans un vecteur physiologiquement acceptable.

La quantité de vecteur qui est injecté à l'organisme hôte choisi varie selon le site de l'injection. A titre indicatif, il peut être injecté entre environ 0,1 et environ 100 µg du polynucléotide codant la protéine ABC1 ans le corps  
25       d'un animal, de préférence d'un patient susceptible de développer une maladie liée à un déficit dans le transport inverse du cholestérol ou ayant déjà développé cette maladie, en particulier un patient ayant une prédisposition pour la maladie de Tangier ou ayant déjà développé la maladie.

30       En conséquence, l'invention concerne également une composition pharmaceutique destinée à la prévention ou au traitement de sujets affectés d'un dysfonctionnement du transport inverse du cholestérol, comprenant un acide nucléique codant pour la protéine ABC1, en association avec un ou plusieurs excipients physiologiquement compatibles.

Avantageusement, une telle composition comprendra le polynucléotide de séquence SEQ ID NO 91, placé sous le contrôle des éléments de régulation appropriés.

L'invention a en outre pour objet une composition pharmaceutique destinée à la prévention ou au traitement de sujets affectés d'un dysfonctionnement du transport inverse du cholestérol, comprenant un vecteur recombinant selon l'invention, en association avec un ou plusieurs excipients physiologiquement compatibles.

L'invention est également relative à l'utilisation d'un acide nucléique selon l'invention, codant pour la protéine ABC1, pour la fabrication d'un médicament destiné à la prévention de l'Athérosclérose sous diverses formes ou plus particulièrement au traitement de sujets affectés d'un dysfonctionnement du transport inverse du cholestérol.

L'invention a également trait à l'utilisation d'un vecteur recombinant selon l'invention, comprenant un acide nucléique codant pour la protéine ABC1, pour la fabrication d'un médicament destiné à la prévention de l'Athérosclérose sous diverses formes ou plus particulièrement au traitement de sujets affectés d'un dysfonctionnement du transport inverse du cholestérol.

20

***Vecteurs utiles dans des procédés de thérapie génique somatique et compositions contenant de tels vecteurs***

La présente invention concerne aussi une nouvelle approche thérapeutique pour le traitement des pathologies liées au transport du cholestérol. Elle propose une solution avantageuse aux inconvénients de l'art antérieur, en démontrant la possibilité de traiter les pathologies liées au transport du cholestérol par la thérapie génique, par le transfert et l'expression in vivo d'un gène codant pour une protéine ABC1 impliquée dans le transport et le métabolisme du cholestérol. L'invention offre ainsi un moyen simple permettant un traitement spécifique et efficace des pathologies associées comme par exemple l'Athérosclérose.

30

La thérapie génique consiste à corriger une déficience ou une anomalie (mutation, expression aberrante, etc) ou à assurer l'expression d'une protéine d'intérêt thérapeutique par introduction d'une information

génétique dans la cellule ou l'organe affecté. Cette information génétique peut être introduite soit ex vivo dans une cellule extraite de l'organe, la cellule modifiée étant alors réintroduite dans l'organisme, soit directement in vivo dans le tissu approprié. Dans ce second cas, différentes techniques existent, parmi lesquelles des techniques diverses de transfection impliquant des complexes d'ADN et de DEAE-dextran (Pagano et al., J.Virol. 1 (1967) 891), d'ADN et de protéines nucléaires (Kaneda et al., Science 243 (1989) 375), d'ADN et de lipides (Felgner et al., PNAS 84 (1987) 7413), l'emploi de liposomes (Fraleigh et al., J.Biol.Chem. 255 (1980) 10431), etc. Plus récemment, l'emploi de virus comme vecteurs pour le transfert de gènes est apparu comme une alternative prometteuse à ces techniques physiques de transfection. A cet égard, différents virus ont été testés pour leur capacité à infecter certaines populations cellulaires. En particulier, les rétrovirus (RSV, HMS, MMS, etc), le virus HSV, les virus adéno-associés, et les adénovirus.

La présente invention est donc également relative à une nouvelle approche thérapeutique pour le traitement des pathologies liées au transport du cholestérol, consistant à transférer et à exprimer in vivo des gènes codant pour ABC1. De manière particulièrement avantageuse, la demanderesse a maintenant montré qu'il est possible de construire des virus recombinants contenant une séquence d'ADN codant pour une protéine ABC1 impliquée dans le métabolisme du cholestérol, d'administrer ces virus recombinants in vivo, et que cette administration permet une expression stable et efficace d'une protéine ABC1 biologiquement active in vivo, et sans effet cytopathologique.

La présente invention résulte également de la mise en évidence que les adénovirus constituent des vecteurs particulièrement efficaces pour le transfert et l'expression du gène ABC1. En particulier, la présente invention montre que l'utilisation d'adénovirus recombinants comme vecteurs permet d'obtenir des niveaux d'expression suffisamment élevés de ce gène pour produire l'effet thérapeutique recherché. D'autres vecteurs viraux tels que les rétrovirus ou les virus adéno-associés (AAV) permettant une expression stable du gène sont aussi revendiqués.

La présente invention offre ainsi une nouvelle approche pour le traitement et la prévention des pathologies cardiovasculaires et neurologiques liées aux anomalies du transport du cholestérol.

L'invention a donc aussi pour objet un virus recombinant défectif comprenant une séquence nucléique codant pour une protéine ABC1 impliquée dans le métabolisme du cholestérol.

L'invention a également trait à l'utilisation d'un tel virus recombinant  
5 défectif pour la préparation d'une composition pharmaceutique destinée au traitement et/ou à la prévention des maladies cardiovasculaires.

La présente invention concerne également l'utilisation de cellules modifiées génétiquement ex vivo par un virus tel que décrit ci-dessus, ou de cellules productrices de tels virus, implantées dans l'organisme, permettant  
10 une expression in vivo prolongée et efficace d'une protéine ABC1 biologiquement active.

La présente invention montre qu'il est possible d'incorporer une séquence d'ADN codant pour ABC1 dans un vecteur viral, et que ces  
15 vecteurs permettent d'exprimer efficacement une forme mature, biologiquement active. Plus particulièrement, l'invention montre que l'expression in vivo de ABC1 peut être obtenue par administration directe d'un adénovirus ou par implantation d'une cellule productrice ou génétiquement modifiée par un adénovirus ou par un rétrovirus incorporant  
20 un tel ADN.

La présente invention est particulièrement avantageuse car elle permet d'induire une expression contrôlée et sans effet nocif de ABC1 dans des organes qui ne sont pas normalement concernés par l'expression de cette protéine. En particulier, une libération significative de la protéine ABC1  
25 est obtenue par implantation de cellules productrices de vecteurs de l'invention, ou infectées ex vivo par des vecteurs de l'invention.

L'activité de transporteur du cholestérol produit dans le cadre de la présente invention peut être du type ABC1 humaine ou animale. La séquence nucléique utilisée dans le cadre de la présente invention peut être  
30 un ADNc, un ADN génomique (ADNg), un ARN (dans le cas des rétrovirus) ou une construction hybride consistant par exemple en un ADNc dans lequel seraient insérés un ou plusieurs introns. Il peut également s'agir de séquences synthétiques ou semi-synthétiques. De manière particulièrement avantageuse, on utilise un ADNc ou un ADNg. En particulier, l'utilisation d'un  
35 ADNg permet une meilleure expression dans les cellules humaines. Pour



permettre leur incorporation dans un vecteur viral selon l'invention, ces séquences sont avantageusement modifiées, par exemple par mutagenèse dirigée, en particulier pour l'insertion de sites de restriction appropriés. Les séquences décrites dans l'art antérieur ne sont en effet pas construites pour  
5 une utilisation selon l'invention, et des adaptations préalables peuvent s'avérer nécessaires, pour obtenir des expressions importantes. Dans le cadre de la présente invention, on préfère utiliser une séquence nucléique codant pour une protéine ABC1 humaine. Par ailleurs, il est également possible d'utiliser une construction codant pour un dérivé de ces protéines  
10 ABC1. Un dérivé de ces protéines ABC1 comprend par exemple toute séquence obtenue par mutation, délétion et/ou addition par rapport à la séquence native, et codant pour un produit conservant l'activité transporteur de cholestérol. Ces modifications peuvent être réalisées par les techniques connues de l'homme du métier (voir techniques générales de biologie  
15 moléculaire ci-après). L'activité biologique des dérivés ainsi obtenus peut ensuite être aisément déterminée, comme indiqué notamment dans les exemples la mesure de l'efflux de cholestérol à partir des cellules. Les dérivés au sens de l'invention peuvent également être obtenus par hybridation à partir de banques d'acides nucléiques, en utilisant comme  
20 sonde la séquence native ou un fragment de celle-ci.

Ces dérivés sont notamment des molécules ayant une plus grande affinité pour leurs sites de fixation, des molécules présentant une plus grande résistance aux protéases, des molécules ayant une efficacité thérapeutique plus grande ou des effets secondaires moindres, ou  
25 éventuellement de nouvelles propriétés biologiques. Les dérivés incluent également les séquences d'ADN modifiées permettant une expression améliorée in vivo.

Dans un premier mode de réalisation, la présente invention concerne un virus recombinant défectif comprenant une séquence d'ADNc codant  
30 pour une protéine ABC1 impliquée dans le transport et le métabolisme du cholestérol. Dans un autre mode de réalisation préféré de l'invention, la séquence d'ADN est une séquence d'ADNg.

Les vecteurs de l'invention peuvent être préparés à partir de différents types de virus. Préférentiellement, on utilise des vecteurs dérivés  
35 des adénovirus, des virus adéno-associés (AAV), des virus de l'herpès

(HSV) ou des rétrovirus. Il est tout particulièrement avantageux d'utiliser un adénovirus, pour une administration directe ou pour la modification ex vivo de cellules destinées à être implantées, ou un rétrovirus, pour l'implantation de cellules productrices.

5 Les virus selon l'invention sont défectifs, c'est-à-dire qu'ils sont incapables de se répliquer de façon autonome dans la cellule cible. Généralement, le génome des virus défectifs utilisés dans le cadre de la présente invention est donc dépourvu au moins des séquences nécessaires à la réplication dudit virus dans la cellule infectée. Ces régions peuvent être  
10 soit éliminées (en tout ou en partie), soit rendues non-fonctionnelles, soit substituées par d'autres séquences et notamment par la séquence nucléique codant pour la protéine ABC1. Préférentiellement, le virus défectif conserve néanmoins les séquences de son génome qui sont nécessaires à l'encapsidation des particules virales.

15 S'agissant plus particulièrement d'adénovirus, différents sérotypes, dont la structure et les propriétés varient quelque peu, ont été caractérisés. Parmi ces sérotypes, on préfère utiliser dans le cadre de la présente invention les adénovirus humains de type 2 ou 5 (Ad 2 ou Ad 5) ou les adénovirus d'origine animale (voir demande WO94/26914). Parmi les  
20 adénovirus d'origine animale utilisables dans le cadre de la présente invention on peut citer les adénovirus d'origine canine, bovine, murine, (exemple : Mav1, Beard et al., Virology 75 (1990) 81), ovine, porcine, aviaire ou encore simienne (exemple : SAV). De préférence, l'adénovirus d'origine animale est un adénovirus canin, plus préférentiellement un adénovirus  
25 CAV2 [souche Manhattan ou A26/61 (ATCC VR-800) par exemple]. De préférence, on utilise dans le cadre de l'invention des adénovirus d'origine humaine ou canine ou mixte. Préférentiellement, les adénovirus défectifs de l'invention comprennent les ITR, une séquence permettant l'encapsidation et la séquence codant pour la protéine ABC1. Avantageusement, dans le  
30 génome des adénovirus de l'invention, la région E1 au moins est rendue non fonctionnelle. Encore plus préférentiellement, dans le génome des adénovirus de l'invention, le gène E1 et au moins l'un des gènes E2, E4, L1-L5 sont non fonctionnels. Le gène viral considéré peut être rendu non fonctionnel par toute technique connue de l'homme du métier, et notamment  
35 par suppression totale, substitution, délétion partielle, ou addition d'une ou

plusieurs bases dans le ou les gènes considérés. De telles modifications peuvent être obtenues in vitro (sur de l'ADN isolé) ou in situ, par exemple, au moyens des techniques du génie génétique, ou encore par traitement au moyen d'agents mutagènes. D'autres régions peuvent également être  
5 modifiées, et notamment la région E3 (WO95/02697), E2 (WO94/28938), E4 (WO94/28152, WO94/12649, WO95/02697) et L5 (WO95/02697). Selon un mode préféré de mise en œuvre, l'adénovirus selon l'invention comprend une délétion dans les régions E1 et E4 et la séquence codant pour ABC1 est insérée au niveau de la région E1 inactivée. Selon un autre mode de  
10 réalisation préféré, il comprend une délétion dans région E1 au niveau de laquelle sont insérés la région E4 et la séquence codant pour ABC1 (Demande de brevet Français FR94 13355).

Les adénovirus recombinants défectifs selon l'invention peuvent être préparés par toute technique connue de l'homme du métier (Leverro et al.,  
15 Gene 101 (1991) 195, EP 185 573; Graham, EMBO J. 3 (1984) 2917). En particulier, ils peuvent être préparés par recombinaison homologue entre un adénovirus et un plasmide portant entre autre la séquence d'ADN codant pour la protéine ABC1. La recombinaison homologue se produit après co-transfection desdits adénovirus et plasmide dans une lignée cellulaire  
20 appropriée. La lignée cellulaire utilisée doit de préférence (i) être transformable par lesdits éléments, et (ii), comporter les séquences capables de compléter la partie du génome de l'adénovirus défectif, de préférence sous forme intégrée pour éviter les risques de recombinaison. A titre d'exemple de lignée, on peut mentionner la lignée de rein embryonnaire  
25 humain 293 (Graham et al., J. Gen. Virol. 36 (1977) 59) qui contient notamment, intégrée dans son génome, la partie gauche du génome d'un adénovirus Ad5 (12 %) ou des lignées capables de compléter les fonctions E1 et E4 telles que décrites notamment dans les demandes n° WO 94/26914 et WO95/02697.

30 Ensuite, les adénovirus qui se sont multipliés sont récupérés et purifiés selon les techniques classiques de biologie moléculaire, comme illustré dans les exemples.

Concernant les virus adéno-associés (AAV), il s'agit de virus à ADN de taille relativement réduite, qui s'intègrent dans le génome des cellules  
35 qu'ils infectent, de manière stable et site-spécifique. Ils sont capables

d'infecter un large spectre de cellules, sans induire d'effet sur la croissance, la morphologie ou la différenciation cellulaires. Par ailleurs, ils ne semblent pas impliqués dans des pathologies chez l'homme. Le génome des AAV a été cloné, séquencé et caractérisé. Il comprend environ 4700 bases, et  
5 contient à chaque extrémité une région répétée inversée (ITR) de 145 bases environ, servant d'origine de réplication pour le virus. Le reste du génome est divisé en 2 régions essentielles portant les fonctions d'encapsidation : la partie gauche du génome, qui contient le gène rep impliqué dans la réplication virale et l'expression des gènes viraux; la partie droite du génome,  
10 qui contient le gène cap codant pour les protéines de capsid du virus.

L'utilisation de vecteurs dérivés des AAV pour le transfert de gènes in vitro et in vivo a été décrite dans la littérature (voir notamment WO 91/18088; WO 93/09239; US 4,797,368, US5,139,941, EP 488 528). Ces demandes décrivent différentes constructions dérivées des AAV, dans  
15 lesquelles les gènes rep et/ou cap sont délétés et remplacés par un gène d'intérêt, et leur utilisation pour transférer in vitro (sur cellules en culture) ou in vivo (directement dans un organisme) ledit gène d'intérêt. Toutefois, aucun de ces documents ne décrit ni ne suggère l'utilisation d'un AAV recombinant pour le transfert et l'expression in vivo ou ex vivo d'une protéine  
20 ABC1, ni les avantages d'un tel transfert. Les AAV recombinants défectifs selon l'invention peuvent être préparés par co-transfection, dans un lignée cellulaire infectée par un virus auxiliaire humain (par exemple un adénovirus), d'un plasmide contenant la séquence codant pour la protéine ABC1 bordée de deux régions répétées inversées (ITR) d'AAV, et d'un  
25 plasmide portant les gènes d'encapsidation (gènes rep et cap) d'AAV. Les AAV recombinants produits sont ensuite purifiés par des techniques classiques.

Concernant les virus de l'herpès et les rétrovirus, la construction de vecteurs recombinants a été largement décrite dans la littérature : voir  
30 notamment Breakfield et al., New Biologist 3 (1991) 203; EP 453242, EP178220, Bernstein et al. Genet. Eng. 7 (1985) 235; McCormick, BioTechnology 3 (1985) 689, etc.

En particulier, les rétrovirus sont des virus intégratifs, infectant les cellules en division. Le génome des rétrovirus comprend essentiellement  
35 deux LTR, une séquence d'encapsidation et trois régions codantes (gag, pol

et env). Dans les vecteurs recombinants dérivés des rétrovirus, les gènes gag, pol et env sont généralement délétés, en tout ou en partie, et remplacés par une séquence d'acide nucléique hétérologue d'intérêt. Ces vecteurs peuvent être réalisés à partir de différents types de rétrovirus tels que notamment le MoMuLV ("murine moloney leukemia virus"; encore désigné MoMLV), le MSV ("murine moloney sarcoma virus"), le HaSV ("harvey sarcoma virus"); le SNV ("spleen necrosis virus"); le RSV ("roux sarcoma virus") ou encore le virus de Friend.

Pour construire des rétrovirus recombinants comportant une séquence codant pour la protéine ABC1 selon l'invention, un plasmide comportant notamment les LTR, la séquence d'encapsidation et ladite séquence codante est généralement construit, puis utilisé pour transfecter une lignée cellulaire dite d'encapsidation, capable d'apporter en trans les fonctions rétrovirales déficientes dans le plasmide. Généralement, les lignées d'encapsidation sont donc capables d'exprimer les gènes gag, pol et env. De telles lignées d'encapsidation ont été décrites dans l'art antérieur, et notamment la lignée PA317 (US4,861,719); la lignée PsiCRIP (WO90/02806) et la lignée GP+envAm-12 (WO89/07150). Par ailleurs, les rétrovirus recombinants peuvent comporter des modifications au niveau des LTR pour supprimer l'activité transcriptionnelle, ainsi que des séquences d'encapsidation étendues, comportant une partie du gène gag (Bender et al., J. Virol. 61 (1987) 1639). Les rétrovirus recombinants produits sont ensuite purifiés par des techniques classiques.

Pour la mise en œuvre de la présente invention, il est tout particulièrement avantageux d'utiliser un adénovirus recombinant défectif. Les résultats donnés ci-après démontrent en effet les propriétés particulièrement intéressantes des adénovirus pour l'expression in vivo d'une protéine ayant une activité de transport de cholestérol. Les vecteurs adénoviraux selon l'invention sont particulièrement avantageux pour une administration directe in vivo d'une suspension purifiée, ou pour la transformation ex vivo de cellules, notamment autologues, en vue de leur implantation. De plus, les vecteurs adénoviraux selon l'invention présentent en outre des avantages importants, tels que notamment leur très haute efficacité d'infection, permettant de réaliser des infections à partir de faibles volumes de suspension virale.

Selon un autre mode particulièrement avantageux de mise en œuvre de l'invention, on utilise une lignée productrice de vecteurs rétroviraux contenant la séquence codant pour la protéine ABC1, pour une implantation in vivo. Les lignées utilisables à cet effet sont notamment les cellules PA317  
5 (US4,861,719), PsiCrip (WO90/02806) et GP+envAm-12 (US5,278,056), modifiées pour permettre la production d'un rétrovirus contenant une séquence nucléique codant pour une protéine ABC1 selon l'invention. Par exemple des cellules souches totipotente, précurseurs des lignées cellulaires sanguines, peuvent être prélevées et isolées chez le sujet. Ces cellules  
10 mises en culture peuvent être alors transfectées par le vecteur rétroviral contenant la séquence codant pour la protéine ABC1 sous le contrôle de promoteurs viraux, non viraux, non viraux et spécifiques des macrophages ou encore sous le contrôle de son propre promoteur. Ces cellules sont alors re-introduites chez le sujet. La différenciation de ces cellules sera à l'origine  
15 de cellules sanguines exprimant la protéine ABC1, notamment à l'origine des monocytes qui, transformés en macrophages, participent à l'élimination du cholestérol de la paroi artériel. Ces macrophages exprimant la protéine ABC1 auront une capacité accrue à métaboliser le cholestérol en excès et le mettront à disposition à la surface cellulaire pour son élimination par les  
20 accepteurs primaires du cholestérol membranaire.

Avantageusement, dans les vecteurs de l'invention, la séquence codant pour la protéine ABC1 est placée sous le contrôle de signaux permettant son expression dans les cellules infectées. Il peut s'agir de signaux d'expression homologues ou hétérologues, c'est-à-dire de signaux  
25 différents de ceux naturellement responsables de l'expression de la protéine ABC1. Il peut s'agir en particulier de séquences responsables de l'expression d'autres protéines, ou de séquences synthétiques. Notamment, il peut s'agir de séquences de gènes eucaryotes ou viraux ou de séquences dérivées, stimulant ou réprimant la transcription d'un gène de façon  
30 spécifique ou non et de façon inductible ou non. A titre d'exemple, il peut s'agir de séquences promotrices issues du génome de la cellule que l'on désire infecter, ou du génome d'un virus, et notamment, les promoteurs des gènes E1A, MLP d'adénovirus, le promoteur CMV, LTR-RSV, etc. Parmi les promoteurs eucaryotes, on peut citer également les promoteurs ubiquitaires  
35 (HPRT, vimentine,  $\alpha$ -actine, tubuline, etc), les promoteurs des filaments

intermédiaires (desmine, neurofilaments, kératine, GFAP, etc) les promoteurs de gènes thérapeutiques (type MDR, CFTR, facteur VIII, etc) les promoteurs spécifiques de tissus (pyruvate kinase, villine, promoteur de la protéine intestinale de liaison des acides gras, promoteur de l'actine  $\alpha$  des  
5 cellules du muscle lisse, promoteurs spécifiques pour le foie ; Apo AI, Apo AII, Albumine humaine etc) ou encore les promoteurs répondant à un stimulus (récepteur des hormones stéroïdes, récepteur de l'acide rétinoïque, etc.). En outre, ces séquences d'expression peuvent être modifiées par addition de séquences d'activation, de régulation, etc. Par ailleurs, lorsque le  
10 gène inséré ne comporte pas de séquences d'expression, il peut être inséré dans le génome du virus défectif en aval d'une telle séquence.

Dans un mode particulier de réalisation, l'invention concerne un virus recombinant défectif comprenant une séquence nucléique codant pour une la protéine ABC1 impliquée dans le métabolisme du cholestérol sous le  
15 contrôle d'un promoteur choisi parmi le LTR-RSV ou le promoteur précoce du CMV.

Comme indiqué ci-avant, la présente invention concerne également toute utilisation d'un virus tel que décrit ci-dessus pour la préparation d'une composition pharmaceutique destinée au traitement et/ou à la prévention  
20 des pathologies liées au transport du cholestérol.

La présente invention concerne également une composition pharmaceutique comprenant un ou plusieurs virus recombinants défectifs tels que décrits précédemment. Ces compositions pharmaceutiques peuvent être formulées en vue d'administrations par voie topique, orale, parentérale,  
25 intranasale, intraveineuse, intramusculaire, sous-cutanée, intraoculaire, transdermique, etc. De préférence, les compositions pharmaceutiques de l'invention contiennent un véhicule pharmaceutiquement acceptable pour une formulation injectable, notamment pour une injection intraveineuse, telle que par exemple dans la veine porte du patient. Il peut s'agir en particulier  
30 de solutions stériles, isotoniques, ou de compositions sèches, notamment lyophilisées, qui, par addition selon le cas d'eau stérilisée ou de sérum physiologique, permettent la constitution de solutés injectables. L'injection directe dans la veine porte du patient est avantageuse car elle permet de cibler l'infection au niveau du foie et ainsi, de concentrer l'effet thérapeutique  
35 au niveau de cet organe.

Les doses de virus recombinant défectif utilisées pour l'injection peuvent être adaptées en fonction de différents paramètres, et notamment en fonction du vecteur viral, du mode d'administration utilisé, de la pathologie concernée ou encore de la durée du traitement recherchée. D'une manière générale, les adénovirus recombinants selon l'invention sont formulés et administrés sous forme de doses comprises entre  $10^4$  et  $10^{14}$  pfu/ml, et de préférence  $10^6$  à  $10^{10}$  pfu/ml. Le terme pfu ("plaque forming unit") correspond au pouvoir infectieux d'une solution de virus, et est déterminé par infection d'une culture cellulaire appropriée, et mesure, généralement après 48 heures, du nombre de plages de cellules infectées. Les techniques de détermination du titre pfu d'une solution virale sont bien documentées dans la littérature.

Concernant les rétrovirus, les compositions selon l'invention peuvent comporter directement les cellules productrices, en vue de leur implantation.

A cet égard, un autre objet de l'invention concerne toute cellule de mammifère infectée par un ou plusieurs virus recombinants défectifs tels que décrits ci-dessus. Plus particulièrement, l'invention concerne toute population de cellules humaines infectée par ces virus. Il peut s'agir en particulier de cellules d'origine sanguine (cellules souche totipotente ou précurseurs), fibroblastes, myoblastes, hépatocytes, kératinocytes, cellules musculaires lisses et endothéliales, cellules gliales, etc.

Les cellules selon l'invention peuvent être issues de cultures primaires. Celles-ci peuvent être prélevées par toute technique connue de l'homme du métier, puis mises en culture dans des conditions permettant leur prolifération. S'agissant plus particulièrement de fibroblastes, ceux-ci peuvent être aisément obtenus à partir de biopsies, par exemple selon la technique décrite par Ham [Methods Cell.Biol. 21a (1980) 255]. Ces cellules peuvent être utilisées directement pour l'infection par les virus, ou conservées, par exemple par congélation, pour l'établissement de banques autologues, en vue d'une utilisation ultérieure. Les cellules selon l'invention peuvent également être des cultures secondaires, obtenues par exemple à partir de banques préétablies (voir par exemple EP 228458, EP 289034, EP 400047, EP 456640).

Les cellules en culture sont ensuite infectées par les virus recombinants, pour leur conférer la capacité de produire une protéine ABC1



biologiquement active. L'infection est réalisée in vitro selon des techniques connues de l'homme du métier. En particulier, selon le type de cellules utilisé et le nombre de copies de virus par cellule désiré, l'homme du métier peut adapter la multiplicité d'infection et éventuellement le nombre de cycles d'infection réalisé. Il est bien entendu que ces étapes doivent être effectuées dans des conditions de stérilité appropriées lorsque les cellules sont destinées à une administration in vivo. Les doses de virus recombinant utilisées pour l'infection des cellules peuvent être adaptées par l'homme du métier selon le but recherché. Les conditions décrites ci avant pour l'administration in vivo peuvent être appliquées à l'infection in vitro. Pour l'infection par des rétrovirus, il est également possible de co-cultiver les cellules que l'on désire infecter avec des cellules productrices des rétrovirus recombinants selon l'invention. Ceci permet de s'affranchir de la purification des rétrovirus.

Un autre objet de l'invention concerne un implant comprenant des cellules de mammifères infectées par un ou plusieurs virus recombinants défectifs telles que décrites ci-dessus ou des cellules productrices de virus recombinants, et une matrice extracellulaire. Préférentiellement, les implants selon l'invention comprennent  $10^5$  à  $10^{10}$  cellules. Plus préférentiellement, ils en comprennent  $10^6$  à  $10^8$ .

Plus particulièrement, dans les implants de l'invention, la matrice extracellulaire comprend un composé gélifiant et éventuellement un support permettant l'ancrage des cellules.

Pour la préparation des implants selon l'invention, différents types de gélifiants peuvent être employés. Les gélifiants sont utilisés pour l'inclusion des cellules dans une matrice ayant la constitution d'un gel, et pour favoriser l'ancrage des cellules sur le support, le cas échéant. Différents agents d'adhésion cellulaire peuvent donc être utilisés comme gélifiants, tels que notamment le collagène, la gélatine, les glycosaminoglycans, la fibronectine, les lectines, etc. De préférence, on utilise dans le cadre de la présente invention du collagène. Il peut s'agir de collagène d'origine humaine, bovine ou murine. Plus préférentiellement, on utilise du collagène de type I.

Comme indiqué ci avant, les compositions selon l'invention comprennent avantageusement un support permettant l'ancrage des

cellules. Le terme ancrage désigne toute forme d'interaction biologique et/ou chimique et/ou physique entraînant l'adhésion et/ou la fixation des cellules sur le support. Par ailleurs, les cellules peuvent soit recouvrir le support utilisé, soit pénétrer à l'intérieur de ce support, soit les deux. On préfère  
5 utiliser dans le cadre de l'invention un support solide, non toxique et/ou bio-compatible. En particulier, on peut utiliser des fibres de polytétrafluoroéthylène (PTFE) ou un support d'origine biologique.

La présente invention offre ainsi un moyen très efficace pour le traitement ou la prévention des pathologies liées au transport du cholestérol, en particulier l'obésité, l'hypertriglycémie, ou, dans le domaine des  
10 affections cardiovasculaires, l'infarctus du myocarde, l'angor, la mort subite, la décompensation cardiaque et les accidents cérébro-vasculaires.

En outre, ce traitement peut concerner aussi bien l'homme que tout animal tel que les ovins, les bovins, les animaux domestiques (chiens, chats,  
15 etc), les chevaux, les poissons, etc.

### CELLULES HOTES RECOMBINANTES

L'invention concerne aussi une cellule hôte recombinante comprenant  
20 l'un quelconque des acides nucléiques de l'invention, et plus particulièrement un acide nucléique de séquence SEQ ID NO 91, 94 ou 96

Selon un autre aspect, l'invention est également relative à une cellule hôte recombinante comprenant un vecteur recombinant tel que ci-dessus décrit.

25 Les cellules hôtes préférées selon l'invention sont par exemple les suivantes :

a) cellules hôtes procaryotes: souches d'*Escherichia coli* (souche DH5- $\alpha$ ), de *Bacillus subtilis*, de *Salmonella typhimurium*, ou encore des  
30 souches d'espèces telles que *Pseudomonas*, *Streptomyces* et *Staphylococcus* ;

b) cellules hôtes eucaryotes: cellules HeLa (ATCC N°CCL2), cellules Cv 1 (ATCC N°CCL70), cellules COS (ATCC N°CRL 1650), cellules Sf-9

(ATCC N°CRL 1711), cellules CHO (ATCC N°CCL-61) ou encore cellules 3T3 (ATCC N°CRL-6361).

### **POLYPEPTIDES ABC1 MUTES**

5 Selon un autre aspect, l'invention concerne un polypeptide codé par un gène ABC1 muté, et plus particulièrement un gène ABC1 muté chez des patients atteints d'un déficit dans le transport inverse du cholestérol, tout particulièrement chez des patients atteints de la maladie de Tangier.

Comme déjà indiqué précédemment, deux mutations délétères ont  
10 été identifiées chez des patients atteints de la maladie de Tangier.

La première mutation correspond à l'insertion d'un fragment d'une centaine de paires de bases au sein de la séquence codante, au niveau de l'exon 12 du gène ABC1, conduisant à la production d'un polypeptide biologiquement inactif de 2233 acides aminés de séquence SEQ ID NO 140.  
15 Le polypeptide ABC1 muté de séquence SEQ ID NO 140 possède, par rapport au polypeptide normal de séquence SEQ ID NO 139, les différences suivantes :

a) une délétion d'un fragment peptidique de séquence "DERKFW" et le remplacement de ce fragment peptidique par la séquence  
20 "EYSGV TSAHCNLCLLSSSDSRASASQVAGITAPATTPG" codée par le fragment nucléotidique de type Alu inséré.

La seconde mutation concerne l'introduction d'un codon stop précoce dans le premier quart de la séquence codante, au niveau de l'exon 13 du gène ABC1, conduisant à la production d'un polypeptide tronqué ayant 574  
25 acides aminés de séquence SEQ ID NO 141. En outre, la délétion de la base G induit un changement du cadre de lecture conduisant à une protéine dont l'extrémité COOH-terminale ne se retrouve pas dans la séquence d'acides aminés du polypeptide ABC1 normal. Il s'agit de la séquence COOH-terminale "RAPRRKLVSICNRCPIPVTLM TSFCG" du polypeptide  
30 ABC1 muté de séquence SEQ ID NO 141.

Ces deux polypeptides sont utiles notamment pour la préparation d'anticorps les reconnaissant spécifiquement. De tels anticorps constituent des moyens de détection de la production de ces polypeptides ABC1 mutés dans un échantillon provenant d'un sujet à tester, préférentiellement d'un  
35 patient présentant des symptômes caractéristiques d'un déficit dans le

transport inverse du cholestérol, et de manière tout à fait préférée chez un patient présentant les symptômes caractéristiques de la maladie de Tangier.

Selon un autre aspect, l'invention concerne donc un polypeptide  
5 comprenant une séquence en acides aminés SEQ ID NO 140.

Selon encore un autre aspect, l'invention concerne un polypeptide comprenant une séquence en acides aminés SEQ ID NO 141.

L'invention est également relative à un polypeptide comprenant une séquence en acides aminés ayant au moins 80% d'identité en acides aminés  
10 avec une séquence en acides aminés choisie dans le groupe constitué des peptides de séquences SEQ ID NO 140 et 141, ou un fragment peptidique de ce dernier.

Un premier fragment peptidique préféré comprendra au moins 5 acides aminés consécutifs du fragment peptidique de séquence  
15 "EYSGVTSAHCNLCLLSSSDSRASASQVAGITAPATTPG" comprise dans le polypeptide ABC1 muté de séquence SEQ ID NO 140.

Un second fragment peptidique préféré comprendra au moins 5 acides aminés consécutifs du fragment peptidique de séquence  
20 "RAPRRKLVSICNRCPIPVTLMTSFCG" comprise dans le polypeptide ABC1 muté de séquence SEQ ID NO 141.

Avantageusement, fait partie de l'invention un polypeptide ayant au moins 85%, 90%, 95% ou 99% d'identité en acides aminés avec une séquence en acides aminés choisie dans le groupe constitué des peptides de séquences SEQ ID NO 140 et 141, ou un fragment peptidique de ce  
25 dernier.

De préférence, des polypeptides selon l'invention auront une longueur de 15, 18 ou 20 à 25, 35, 40, 50, 70, 80, 100 ou 200 acides aminés consécutifs d'un acide nucléique selon l'invention, en particulier un  
30 polypeptide de séquence en acides aminés choisie parmi les séquences SEQ ID N°140 et 141.

Alternativement, un polypeptide selon l'invention consistera et/ou comprendra les fragments d'une longueur de 15, 18, 20, 25, 35, 40, 50, 100  
35 ou 200 acides aminés consécutifs d'un polypeptide selon l'invention, plus

particulièrement d'un polypeptide choisi parmi les séquences SEQ ID NO 140 et 141.

De manière générale, les polypeptides selon la présente invention se présentent sous une forme isolée ou purifiée.

L'invention est également relative à un procédé pour la production de l'un des polypeptides de séquences SEQ ID NO 140 et 141, ou d'un fragment peptidique ou d'un variant de ce dernier, ladite méthode comprenant les étapes de :

- a) insérer un acide nucléique codant pour ledit polypeptide dans un vecteur approprié ;
- b) cultiver, dans un milieu de culture approprié, une cellule hôte préalablement transformée ou transfecter avec le vecteur recombinant de l'étape a) ;
- c) récupérer le milieu de culture conditionné ou lyser la cellule hôte, par exemple par sonication ou par choc osmotique ;
- d) séparer et purifier à partir dudit milieu de culture ou encore à partir des lysats cellulaires obtenus à l'étape c), ledit polypeptide ;
- e) le cas échéant, caractériser le polypeptide recombinant produit.

Les peptides selon l'invention peuvent être caractérisés par fixation sur une colonne de chromatographie d'immunoaffinité sur laquelle les anticorps dirigés contre ce polypeptide ou contre un fragment ou un variant de ce dernier ont été préalablement immobilisés.

Selon un autre aspect, un polypeptide recombinant selon l'invention peut être purifié par passage sur une série appropriée de colonnes de chromatographie, selon les méthodes connues de l'homme de l'art et décrites par exemple dans F.Ausubel et al (1989).

Un polypeptide selon l'invention peut être également préparé par les techniques classiques de synthèse chimique indifféremment en solution homogène ou phase solide.

A titre illustratif, un polypeptide selon l'invention pourra être préparé par la technique ou en solution homogène décrite par Houben Weyl (1974)

ou encore la technique de synthèse en phase solide décrite par Merrifield (1965a; 1965b).

Font également partie de l'invention des polypeptides dits "homologues" à l'un quelconque des polypeptides de séquences d'acides aminés SEQ ID NO 140 et 141, ou de leurs fragments ou variants.

De tels polypeptides homologues ont des séquences d'acides aminés possédant une ou plusieurs substitutions d'un acide aminé par un acide aminé équivalent, par rapport aux polypeptides de référence.

On entendra par acide aminé équivalent selon la présente invention, par exemple remplacement d'un résidu sous la forme L par un résidu sous la forme D ou encore le remplacement d'un acide glutamique (E) par un acide pyro-glutamique selon des techniques bien connues de l'homme du métier. A titre illustratif, la synthèse de peptide contenant au moins un résidu sous la forme D est décrite par Koch (1977).

Selon un autre aspect, sont également considérés comme des acides aminés équivalents deux acides aminés appartenant à la même classe, c'est-à-dire deux acides aminés acide, basique, non polaire ou encore polaire non chargé.

Font également partis de l'invention des polypeptides comprenant au moins une liaison non peptidique telle qu'une liaison rétro-inverso (NHCO), une liaison carba ( $\text{CH}_2\text{CH}_2$ ) ou encore une liaison cétométhylène ( $\text{CO}-\text{CH}_2$ ).

De préférence, les polypeptides selon l'invention comprenant une ou plusieurs additions, délétions, substitutions d'au moins un acide aminé conserveront leur capacité à être reconnus par des anticorps dirigés contre les polypeptides non modifiés.

## ANTICORPS

Les polypeptides ABC1 mutés selon l'invention, en particulier les polypeptides de séquences en acides aminés SEQ ID NO 140-141] ou les fragments de ces derniers ainsi que les peptides homologues peuvent être utilisés pour la préparation d'anticorps, notamment en vue de détecter la production de formes altérés du polypeptide ABC1 chez un patient.

Un premier anticorps préféré selon l'invention est dirigé contre un fragment peptidique de séquence

"EYSGVTSAHCNLCLLSSSDSRASASQVAGITAPATTPG" comprise dans le polypeptide ABC1 muté de séquence SEQ ID NO 140.

Un second anticorps préféré selon l'invention est dirigé contre un fragment peptidique comprenant au moins 5 acides aminés consécutifs du  
5 fragment peptidique de séquence "RAPRRKLVSICNRCPIPVTLMTSFCG" comprise dans le polypeptide ABC1 muté de séquence SEQ ID NO 141.

Par "anticorps" au sens de la présente invention, on entendra notamment des anticorps polyclonaux ou monoclonaux ou des fragments (par exemple les fragments F (ab)<sub>2</sub>, Fab) ou encore tout polypeptide  
10 comprenant un domaine de l'anticorps initial reconnaissant le polypeptide ou le fragment de polypeptide cible selon l'invention.

Des anticorps monoclonaux peuvent être préparés à partir d'hybridomes selon la technique décrite par Kohler et Milstein (1975).

15 La présente invention concerne également des anticorps dirigés contre un polypeptide tel que décrit ci-dessus ou un fragment ou un variant de ce dernier, tels que produits dans la technique du trioma ou encore la technique d'hybridome décrite par Kozbor et al. (1983).

20 L'invention a également trait à des fragments d'anticorps simple chaîne Fv (ScFv) tels que décrits dans le brevet US N° 4,946,778 ou encore par Martineau et al. (1998).

Les anticorps selon l'invention comprennent également des fragments d'anticorps obtenus à l'aide de banques de phages Ridder et al., (1995) ou  
25 encore des anticorps humanisés Reinmann et al. (1997); Leger et al., (1997).

Les préparations d'anticorps selon l'invention sont utiles dans des tests de détection immunologiques destinés à l'identification de la présence et/ou de la quantité d'antigènes présents dans un échantillon.

30 Un anticorps selon l'invention pourra comprendre en outre un marqueur détectable isotopique ou non-isotopique, par exemple fluorescent ou encore être couplé à une molécule telle que la biotine, selon des techniques bien connues de l'homme du métier.

Ainsi, la mention a en outre pour objet un procédé pour détecter la présence d'un polypeptide conforme à l'invention dans un échantillon, ledit procédé comprenant les étapes de :

- a) mettre en contact l'échantillon à tester avec un anticorps tel que décrit ci-dessus ;
- b) détecter le complexe antigène/anticorps formé.

L'invention est également relative à un nécessaire ou kit de diagnostic ou pour la détection de la présence d'un polypeptide conforme à l'invention dans un échantillon, ledit nécessaire comprenant :

- a) un anticorps tel que défini ci-dessus ;
- b) un réactif permettant la détection des complexes antigène/anticorps formés.

## **COMPOSITIONS PHARMACEUTIQUES ET METHODES DE TRAITEMENT THERAPEUTIQUES**

L'invention concerne aussi des compositions pharmaceutiques destinées à la prévention ou au traitement d'un déficit dans le métabolisme du cholestérol telle que l'athérosclérose, particulièrement dans le transport du cholestérol, et plus particulièrement encore dans le transport inverse du cholestérol, caractérisées en ce qu'elles comprennent une quantité thérapeutiquement efficace d'un polynucléotide capable de donner lieu à la production d'une quantité efficace du polypeptide ABC1 normal, en particulier du polypeptide de séquence SEQ ID NO 139.

L'invention a en outre pour objet des compositions pharmaceutiques destinées à la prévention ou au traitement d'un déficit dans le métabolisme du cholestérol telle que l'athérosclérose, particulièrement dans le transport du cholestérol, et plus particulièrement encore dans le transport inverse du cholestérol, caractérisées en ce qu'elles comprennent une quantité thérapeutiquement efficace du polypeptide ABC1 normal, en particulier du polypeptide de séquence SEQ ID NO 139.

De telles compositions pharmaceutiques seront avantageusement adaptées pour l'administration, par exemple par voie parentérale, d'une quantité du polypeptide ABC1 allant de 1 µg/kg/jour à 10 mg/kg/jour, de



préférence au moins 0,01 mg/kg/jour et de manière tout à fait préférée entre 0,01 et 1 mg/kg/jour.

Les compositions pharmaceutiques selon l'invention peuvent être indifféremment administrées par voie orale, rectale, parentérale, intra-veineuse, sous-cutanée ou encore intra-dermique.

L'invention concerne aussi l'utilisation du polypeptide ABC1 de séquence SEQ ID NO 139 pour la fabrication d'un médicament destiné à la prévention de l'Athérosclérose sous diverses formes ou plus particulièrement au traitement de sujets affectés d'un dysfonctionnement du transport inverse du cholestérol.

L'invention est enfin relative à une composition pharmaceutique pour la prévention ou le traitement de sujets affectés d'un dysfonctionnement du transport inverse du cholestérol, comprenant une quantité thérapeutiquement efficace du polypeptide de séquence SEQ ID NO 139

Selon un autre aspect, l'invention a également pour objet une méthode de traitement thérapeutique préventive ou curative de maladies provoquées par une déficience dans le métabolisme du cholestérol, plus particulièrement dans le transport du cholestérol et encore plus particulièrement dans le transport inverse du cholestérol, une telle méthode comprenant une étape au cours de laquelle est administrée à un patient un polynucléotide capable de donner lieu à l'expression du polypeptide ABC1 chez ledit patient, ledit polynucléotide étant, le cas échéant, associé à un ou plusieurs véhicules et/ou excipients physiologiquement compatibles.

De préférence, on administrera au patient une composition pharmaceutique comprenant un polynucléotide, telle que définie ci-dessus.

Selon encore un autre aspect, l'invention a également pour objet une méthode de traitement thérapeutique préventive ou curative de maladies provoquées par une déficience dans le métabolisme du cholestérol, plus particulièrement dans le transport du cholestérol et encore plus particulièrement dans le transport inverse du cholestérol, une telle méthode comprenant une étape au cours de laquelle est administrée à un patient une quantité thérapeutiquement efficace du polypeptide ABC1 chez ledit patient,

ledit polypeptide étant, le cas échéant, associé à un ou plusieurs véhicules et/ou excipients physiologiquement compatibles.

De préférence, on administrera au patient une composition pharmaceutique comprenant un polypeptide, telle que définie ci-dessus.

5

#### **METHODES DE CRIBLAGE D'UN COMPOSE AGONISTE OU ANTAGONISTE DU POLYPEPTIDE ABC1.**

Selon un autre aspect, l'invention concerne aussi divers procédés de criblage de composés à visée thérapeutique utiles dans le traitement de  
10 maladies dues à un déficit dans le métabolisme du cholestérol, particulièrement dans le transport du cholestérol, plus particulièrement encore dans le transport inverse du cholestérol, telles que la maladie de Tangier, ou plus généralement les affections de type FHD.

L'invention est donc également relative à l'utilisation du polypeptide  
15 ABC1, ou de cellules exprimant le polypeptide ABC1, pour cribler des principes actifs pour la prévention ou le traitement de maladies résultant d'un dysfonctionnement du transport inverse du cholestérol.

Les sites catalytiques et fragments oligopeptidiques ou immunogéniques du polypeptide ABC1 peuvent servir pour cribler des banques de produits par  
20 tout une foule de techniques existantes. Le fragment utilisé dans ce type de criblage peut être libre en solution, fixé sur un support solide, à la surface cellulaire ou encore dans la cellule. La formation des complexes de liaisons entre les fragments d'ABC1 et l'agent testé peut alors être mesuré.

25 Une autre technique de criblage de produit qui peut être utilisée dans les criblages à haut flux donnant accès à des produits ayant de l'affinité pour la protéine d'intérêt est décrite dans l'application WO84/03564. Dans cette méthode, appliquée à la protéine ABC1, différents produits sont synthétisés sur une surface solide. Ces produits réagissent avec la protéine ABC1 ou  
30 des fragments de celle-ci et le complexe est lavé. Les produits liant la protéine ABC1 sont ensuite détectés par des méthodologies connues de l'homme de l'art. Des anticorps non neutralisants peuvent aussi être utilisés pour capturer un peptide et l'immobiliser sur un support.

Une autre possibilité est d'utiliser un criblage de produit utilisant la compétition d'anticorps neutralisants ABC1, la protéine ABC1 et un produit potentiellement liant la protéine ABC1. De cette manière, les anticorps peuvent être utilisés pour détecter la présence de peptide ayant des motifs  
5 antigéniques commun avec ABC1.

Dans les produits à évaluer et permettant d'augmenter l'activité d'ABC1, on mentionnera notamment les homologues d'ATP spécifiques des kinases impliquées dans l'activation de la molécules ainsi que des phosphatases qui  
10 pourront éviter la déphosphorylation issue de ces dites kinases. On nommera notamment les inhibiteurs de du type phosphodiesterase (PDE) théophylline et 3-isobutyl-1-méthylxanthine ou les activateurs d'adénylcyclase forskoline.

15 Aussi, nous revendiquons dans cette invention l'utilisation de tout procédé de criblage de produits basé sur la méthode de translocation du cholestérol (voir l'exemple 17) entre les membranes ou vésicules, et ceci dans tous les types synthétiques ou cellulaires, c'est à dire de mammifères, d'insectes, de bactéries ou de levures exprimant de manière constitutive ou bien ayant  
20 incorporés la séquence d'ABC1 humaine. A cet effet, des analogues lipidiques marqués peuvent être utilisés.

De même, il a été décrit que la protéine ABC1 permettait le transport d'anion (Becq et al. Journal of Biological Chemistry vol 272, n°5 pages 2695-2699,  
25 1997 et Yamon et al. Blood vol 90, n°8 pages 2911-2915, 1997) et ce transport était activé par des inhibiteurs de phosphatase comme l'acide okadaïque et l'orthovanadate ainsi que par l'élévation de l'AMPc par des agents comme la forskoline. Nous revendiquons l'utilisation de ce système pour cribler des molécules modulant l'activité de la protéine ABC1 (voir  
30 Exemple 18).

Yamon et al (Blood vol 90, n°8 pages 2911-2915, 1997) ont démontré que la protéine ABC1 de souris était impliquée dans la sécrétion d'une cytokine pro-inflammatoire IL-1 $\beta$  dans des macrophages péritoneaux de souris. On  
35 peut donc aussi proposer une méthode de criblage de produits modulant

l'activité de la protéine ABC1 par détermination du relargage de l'IL-1beta à partir de tout type cellulaire exprimant deux protéines (voir Exemple 19).

De plus, sachant que la disruption de nombreux transporteurs ont été décrits  
5 (van, den Hazel. H., H. Pichler, V. M. M. do, E. Leitner, A. Goffeau, and G. Daum. 1999. PDR16 and PDR17, two homologous genes of *Saccharomyces cerevisiae*, affect lipid biosynthesis and resistance to multiple drugs. *J Biol Chem.* 274 (4):1934-41), on peut penser utiliser des mutants cellulaires ayant un phénotype caractéristique et compléter la fonction de ceux-ci  
10 par ABC1 et utiliser l'ensemble à des fins de criblage.

L'invention est également relative à un procédé de criblage d'un composé actif sur le métabolisme du cholestérol, agoniste ou antagoniste du polypeptide ABC1, ledit procédé comprenant les étapes suivantes :

- 15 a) préparer des vésicules membranaires contenant le polypeptide ABC1 et un substrat lipidique comprenant un marqueur détectable ;
- b) incuber les vésicules obtenues à l'étape a) avec un composé candidat agoniste ou antagoniste ;
- c) mesurer qualitativement et/ou quantitativement la libération du substrat  
20 lipidique comprenant un marqueur détectable ;
- d) comparer la mesure obtenue à l'étape b) avec une mesure de la libération du substrat lipidique marqué par les vésicules n'ayant pas préalablement été incubées avec le composé candidat agoniste ou antagoniste.

25 Selon un premier aspect du procédé de criblage ci-dessus, les vésicules membranaires sont des vésicules lipidiques synthétiques, qui peuvent être préparées selon des techniques bien connues de l'homme du métier. Selon cet aspect particulier, la protéine ABC1 peut être une protéine ABC1 recombinante.

30

Selon un second aspect, les vésicules membranaires sont des vésicules de membranes plasmiques dérivées de cellules exprimant le polypeptide ABC1. Il peut s'agir de cellules exprimant naturellement le polypeptide ABC1 ou encore de cellules transfectées avec un vecteur recombinant codant pour le  
35 polypeptide ABC1.

Selon un troisième aspect du procédé de criblage ci-dessus, le substrat lipidique est choisi parmi le cholestérol ou la phosphatidyl choline.

- 5 Selon un quatrième aspect, le substrat lipidique est marqué radioactivement, par exemple par un isotope choisi parmi le  $^3\text{H}$  ou  $^{125}\text{I}$ .

Selon un cinquième aspect, le substrat lipidique est marqué par un composé fluorescent, tel que le NBD ou le pyrène.

10

Selon un sixième aspect, les vésicules membranaires comprenant le substrat lipidique marqué et le polypeptide ABC1 sont immobilisées à la surface d'un support solide préalablement à l'étape b).

- 15 Selon un septième aspect, la mesure de la fluorescence ou de la radioactivité libérée par les vésicules est le reflet direct de l'activité de transport du substrat lipidique par le polypeptide ABC1.

20 L'invention est encore relative à un procédé de criblage d'un composé actif sur le métabolisme du cholestérol, agoniste ou antagoniste du polypeptide ABC1, ledit procédé comprenant les étapes suivantes :

- a) obtenir des cellules, par exemple une lignée cellulaire, exprimant naturellement ou après transfection le polypeptide ABC1 ;  
b) incuber les cellules de l'étape a) en présence d'anion marqué par un  
25 marqueur détectable ;  
c) laver les cellules de l'étape b) afin d'éliminer l'excès de l'anion marqué n'ayant pas pénétré dans ces cellules ;  
d) incuber les cellules obtenues à l'étape c) avec un composé candidat agoniste ou antagoniste du polypeptide ABC1 ;  
30 e) mesurer l'efflux de l'anion marqué ;  
f) comparer la valeur de l'efflux de l'anion marqué déterminé à l'étape e) avec la valeur de l'efflux de l'anion marqué mesuré avec des cellules n'ayant pas préalablement été incubées en présence du composé candidat agoniste ou antagoniste du polypeptide ABC1.

35

Selon un premier aspect du procédé de criblage ci-dessus, les cellules mises en œuvre sont des cellules exprimant naturellement le polypeptide ABC1. Il peut s'agir de monocytes humains en culture primaire, purifiés à partir d'une population de cellules mononucléées du sang humain. Il peut s'agir aussi de  
5 lignées de cellules monocytaires humaines, telles que la lignée leucémique monocyttaire THP1.

Selon un second aspect, les cellules mises en œuvre dans le procédé de criblage décrit ci-dessus peuvent être des cellules n'exprimant pas  
10 naturellement, ou alternativement exprimant à un faible niveau, le polypeptide ABC1, lesdites cellules étant transfectées avec un vecteur recombinant selon l'invention capable de diriger l'expression du polypeptide ABC1

15 Selon un troisième aspect, les cellules peuvent être des cellules ayant un déficit naturel dans le transport anionique, ou encore des cellules pré-traitées avec un ou plusieurs inhibiteurs des canaux anioniques, tels que le Verapamil™ ou le tétra éthylammonium.

20 Selon un quatrième aspect dudit procédé de criblage, l'anion est un iodure marqué radioactivement, tels que les sels  $K^{125}I$  ou  $Na^{125}I$ .

Selon un cinquième aspect, la mesure de l'efflux de l'anion marqué est déterminé périodiquement au cours du temps pendant l'expérience,  
25 permettant ainsi d'établir aussi une mesure cinétique de cet efflux.

Selon un sixième aspect, la valeur de l'efflux de l'anion marqué est déterminée par la mesure de la quantité de l'anion marqué présent à un temps donné dans le surnageant de culture cellulaire.

30 Selon un septième aspect, la valeur de l'efflux de l'anion marqué est déterminée comme la proportion de radioactivité retrouvée dans le surnageant de culture cellulaire par rapport à la radioactivité totale correspondant à la somme de la radioactivité retrouvée dans les lysats

cellulaires et de la radioactivité retrouvée dans le surnageant de culture cellulaire.

L'invention a encore pour objet un procédé de criblage d'un composé actif  
5 sur le métabolisme du cholestérol, agoniste ou antagoniste du polypeptide ABC1, ledit procédé comprenant les étapes suivantes :

- a) cultiver des cellules d'une lignée monocyttaire humaine dans un milieu de culture approprié, en présence d'albumine humaine purifiée ;
- b) incuber les cellules de l'étape a) simultanément en présence d'un  
10 composé stimulant la production d'IL-1 beta et du composé candidat agoniste ou antagoniste;
- c) incuber les cellules obtenues à l'étape b) en présence d'une concentration appropriée d'ATP ;
- d) mesurer d'IL-1 beta libérée dans le surnageant de culture cellulaire.
- 15 e) comparer la valeur de la libération de l'IL-1 beta obtenue à l'étape d) à la valeur de l'IL-1 beta libérée dans le surnageant de culture de cellules n'ayant pas été préalablement incubées en présence du composé candidat agoniste ou antagoniste.

- 20 Selon un premier aspect du procédé de criblage décrit ci-dessus, les cellules utilisées appartiennent à la lignée monocyttaire leucémique humaine THP1.

Selon un second aspect du procédé criblage, le composé stimulant la production d'IL-1 beta est un lipopolysaccharide.

25

Selon un troisième aspect dudit procédé, on détermine également qualitativement et/ou quantitativement la production d'IL-1 alpha, de IL-6 et de TNF alpha par ces cellules.

- 30 Selon un quatrième aspect, le niveau d'expression de l'ARN messager codant pour l'IL-1 beta est également déterminé.

L'invention est illustrée, sans pour autant être limitée, par les figures et les exemples suivants :

La **Figure 1** illustre la ségrégation de la mutation par insertion d'une séquence Alu dans l'exon 12 du gène ABC1. L'insertion-délétion dans l'exon 12 du gène ABC1 consiste en une délétion de 14 nucléotides et d'une insertion de 110 nucléotides comme représenté sur la figure 1A. La figure 1B  
5 représente le pédigré Nu et la taille des fragments d'ADN obtenus pour chacun des patients après amplification par PCR de l'exon 12. La piste M correspond aux marqueurs de mobilité (Gibco BRL) . La piste C correspond à un ADN témoin.

10 La **Figure 2** illustre la mutation par délétion d'un seul nucléotide dans l'exon 13 du gène ABC1. La séquence du brin complémentaire a été obtenue et est représentée de 3' vers 5'. La séquence codant pour les acides aminés 546 à 552 du polypeptide ABC1 est représentée pour différents patients de la famille étudiée, respectivement pour un individu homozygote (**Figure 2-a**),  
15 hétérozygote (**Figure 2-b**) et non affecté (**Figure 2-c**). La séquence de la **Figure 2-a** a été retrouvée chez trois individus homozygotes, la séquence de la **Figure 2-b** a été retrouvée chez cinq individus hétérozygotes et la séquence de la **Figure 2-c** a été retrouvée chez quatre individus non affectés.

20

## EXEMPLES

### EXEMPLE 1: Distribution tissulaire des transcrits du gène ABC1 selon l'invention

25 Le profil d'expression des polynucleotides selon la présente invention est déterminé selon les protocoles d'analyse de Northern blot et de transcription inverse couplée à la PCR décrits notamment par Sambrook et al (ref. CSH Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). "Molecular Cloning : A Laboratory Manual," 2<sup>nd</sup> ed., Cold Spring Harbor Laboratory Press, Cold  
30 Spring Harbor, NY.).

Par exemple, dans le cas d'une analyse par transcription inverse, un couple d'amorces synthétisées à partir de l'ADN complet du gène humain ABC1 de séquence SEQ ID NO 91 est utilisé pour détecter l'ADNc  
35 correspondant.



La réaction de polymérase en chaîne (PCR) est réalisée sur des matrices d'ADNc correspondant à des ARNm polyA<sup>+</sup> (Clontech) rétrotranscrits. La transcription inverse en ADNc est réalisée avec l'enzyme SUPERScript II (GibcoBRL, Life Technologies) selon les conditions décrites par le fabricant.

La réaction de polymérase en chaîne est réalisée selon des conditions standard, dans 20 µl de mélange réactionnel avec 25 ng de la préparation d'ADNc. Le mélange réactionnel est composé de 400 µM de chacun des dNTP, de 2 unités de *Thermus aquaticus* (Taq) ADN polymérase (Ampli Taq Gold ; Perkin Elmer), de 0,5 µM de chaque amorce, de 2,5 mM MgCl<sub>2</sub>, et de tampon PCR. Trente quatre cycles de PCR (dénaturation 30 s à 94 °C, hybridation de 30 s décomposé comme suit lors des 34 cycles : 64°C 2 cycles, 61°C 2 cycles, 58°C 2 cycles et 55°C 28 cycles et une élongation d'une minute par kilobase à 72°C) sont réalisés après une première étape de dénaturation à 94°C durant 10 min dans un thermocycler Perkin Elmer 9700. Les réactions de PCR sont visualisées sur gel d'agarose par électrophorèse. Les fragments d'ADNc obtenus peuvent être utilisés comme sondes pour une analyse par Northern blot et peuvent également être utilisés pour la détermination exacte de la séquence polynucléotidique.

20

Dans le cas d'une analyse par Northern Blot, une sonde d'ADNc produite comme décrit ci-dessus est marquée au <sup>32</sup>P grâce au système de marquage d'ADN High Prime (Boehringer) selon les instructions indiquées par le fabricant. Après marquage, la sonde est purifiée sur une microcolonne de Sephadex G50 (Pharmacia) selon les instructions indiquées par le fabricant. La sonde marquée et purifiée est alors utilisée pour la détection de l'expression des ARNm dans différents tissus.

Le Northern blot contenant des échantillons d'ARN de différents tissus humains ((Multiple Tissue Northern , MTN, Clontech) Blot 2, référence 77759-1) est hybridé avec la sonde marquée.

Le protocole suivi pour les hybridations et lavages peut être soit directement celui décrit par le fabricant (Manuel d'utilisation PT1200-1) soit une adaptation de ce protocole en utilisant les méthodes connues de l'homme de l'art et décrites par exemple dans F.Ausubel et al (1999). On

pourra ainsi faire varier par exemple les températures de préhybridation et d'hybridation en présence de formamide.

Par exemple on pourra utiliser le protocole suivant :

1- Compétition des membranes et PRE HYBRIDATION :

5

- Mélanger : 40µl ADN sperme de saumon (10mg/ml)  
+ 40 µl ADN placentaire humain (10mg/ml)
- Dénaturer 5 mn à 96°C, puis plonger dans la glace le mélange.

10

- Oter le SSC 2X et verser 4 ml de mix formamide dans le tube d'hybridation contenant les membranes.

- Ajouter le mélange des deux ADN dénaturés.

15

- Incubation à 42°C pendant 5 à 6 heures, avec rotation.

2- Compétition de la sonde marquée :

- 20 - Ajouter à la sonde marquée et purifiée 10 à 50 µl ADN Cot I, selon la quantité de séquences répétées.

- Dénaturer 7 à 10 mn à 95°C.

- 25 - Incuber à 65°C pendant 2 à 5 heures.

3- HYBRIDATION :

- Oter le mix de pré-hybridation.

30

- Mélanger 40 µl ADN sperme de saumon + 40 µl ADN placentaire humain ;  
dénaturer 5 mn à 96°C, puis plonger dans la glace.

- Ajouter dans le tube d'hybridation 4 ml de mix formamide, le mélange des  
35 deux ADN et la sonde marquée/ADN Cot I dénaturée.

- Incuber 15 à 20 heures à 42°C, avec rotation.

#### 4- Lavages :

5

- Un lavage à température ambiante dans du SSC 2X, pour rincer.
- 2 fois 5 minutes à température ambiante SSC 2X et SDS 0,1% à 65°C.
- 2 fois 15 minutes à 65°C SSC 1X et SDS 0,1% à 65°C.

10

Après hybridation et lavage, le blot est analysé après une nuit d'exposition au contact d'un écran au phosphore révélé à l'aide du Storm (Molecular Dynamics, Sunnyvale, CA).

#### 15 **EXEMPLE 2 : Obtention de l'ADNc complet du gène ABC1**

La séquence de la région 3'-UTR de l'ADNc du gène ABC1 humain a été identifiée par recherche dans les bases de données.

Un criblage itératif d'une base de données de séquences EST ( "Genbank mouse human subdivision EST, v.111") a été réalisé à l'aide du logiciel

20 **BLAST.**

Des amorces oligonucléotidiques ont été synthétisées à partir de la séquence consensus partielle dérivée des séquences EST, afin d'amplifier par une réaction de RT-PCR l'extrémité 3' de l'ADNc du gène ABC1 humain, puis d'en déterminer la séquence.

25 Les amorces oligonucléotidiques utilisées sont les suivantes :

1. 5'- AAACCAGACAGTAGTGGACG-3', (SEQ ID NO 142)
2. 5'-GTTACTGCCACCAGAACAGC-3', (SEQ ID NO 143)
3. 5'- TGATAAGCTGTTCTGGTGGC-3',(SEQ ID NO 144)
4. 5'-CTTGGCTTTTGCATTGTTGC-3', (SEQ ID NO 145)
- 30 5. 5'- CAATGCAAAAGCCAAGAAAG-3',(SEQ ID NO 146)
6. 5'-TGCAACGATGCCATATCAC-3', (SEQ ID NO 147)
7. 5'-CAACTCCTTACTTCGGTTCCTC-3',(SEQ ID NO 148)
8. 5'-GTTTTCTGAGGTGTCCCAAAG-3'( SEQ ID NO 149)

La transcription inverse de l'ARNm poly(A)+ à partir de tissus de

35 

cerveau, de cerveau foetal, de cœur, d'utérus et de placenta (Banques

commercialisées par la société Clontech) a été réalisé par élongation à l'aide d'amorces oligodT en utilisant le kit Superscript™ (commercialisé par la société Life Technologies Inc.), selon les instructions du fabricant.

Dans chaque expérience, on a pu exclure la présence d'ADN contaminant du fait de l'absence de polynucléotides amplifiés par PCR dans les échantillons ne contenant pas de transcriptase inverse.

Une réaction PCR a été effectuée sur les produits ayant subi ou non une première étape de transcription inverse dans les conditions suivantes :

- 400 µM dNTP, 2 Unités d'ADN polymérase Taq (*Thermus aquaticus*, Ampli Taq Gold, commercialisée par la société Perkin Elmer), 0,5 µM de chacune des amorces, 2,5 mM de Mg Cl<sub>2</sub>, l'ensemble étant présent dans un tampon pour PCR contenant également 50 ng d'ADN ou environ 25 ng d'ADNc.
- La réaction PCR a été effectuée pendant 30 cycles dans un appareil thermocycleur ("Perkin Elmer 9700 Thermal Cycler") dans des micro-plaques de 96 puits.
- Après une dénaturation initiale à 94°C pendant 10 minutes, chaque cycle a été réalisé de la manière suivante :
  - étape de dénaturation à 94°C pendant 30 secondes; étape d'hybridation pendant 30 secondes (à 64°C pendant 2 cycles, à 61°C pendant 2 cycles, à 58°C pendant 2 cycles et à 55°C pendant 28 cycles).
  - étape d'élongation pendant un temps correspondant à 1 minute par kilobase.
- La réaction PCR est arrêtée par une étape finale d'extension de 7 minutes.

Différentes autres approches peuvent être utilisées pour isoler l'ADNc correspondant à l'ADNc complet de ABC1.

Par exemple un clone complet peut être directement isolé par hybridation en criblant une banque d'ADNc au moyen d'une sonde polynucléotidique spécifique de la séquence du gène d'intérêt.

En particulier une sonde spécifique de 30-40 nucléotides est synthétisée en utilisant un synthétiseur de marque Applied Biosystem/Perkin Elmer selon la séquence choisie.

L'oligonucléotide obtenu est radiomarké, par exemple au <sup>32</sup>P-γ- ATP en utilisant la T4 polynucléotide kinase et est purifié selon les méthodes usuelles (e.g. Maniatis et al. Molecular cloning : A Laboratory Manual, Cold

Spring Harbor Press, Cold Spring, NY 1982 ou encore F.Ausubel et al .  
(Current Protocols in Molecular Biology, J.Wiley and Sons Eds, 1989).

La banque de clones contenant l'ADNc que l'on veut cribler est étalée sur  
5 milieu de culture en boîte de Pétri (1.5% agar) contenant les antibiotiques  
appropriés selon les méthodes usuelles citées ci dessus (F.Ausubel et al.).  
Les colonies ainsi produites après incubation sont transférées sur filtres de  
nitrocellulose et criblées au moyen de la sonde nucléotidique radiomarquée,  
selon les méthodes usuelles et les colonies hybridant avec la sonde sont  
10 isolées et sous clonées.

L'ADN des clones ainsi repéré est préparé et analysé par séquençage. Les  
clones contenant les fragments correspondant à l'ADNc complet sont  
purifiés et reclés dans le vecteur pcDNA3 selon les protocoles connus de  
15 l'homme de l'art et présentés par exemple dans F. Ausubel et al (1989) .

Différentes méthodes sont connues pour identifier les extrémités 5' et 3' du  
cDNA correspondant aux gènes décrits dans la présente demande. Ces  
méthodes incluent mais ne se limitent pas au clonage par hybridation, au  
20 clonage utilisant des protocoles similaires ou identiques à la 3' ou 5' RACE-  
PCR (Rapid Amplification of cDNA End-PCR) qui sont bien connues de  
l'homme de l'art.

Par exemple, on pourra utiliser le kit commercialisé par la société Clontech  
25 (Marathon Ready™ cDNA kit , protocole référencé PT1156-1) ou  
alternativement une méthode similaire à la 5'RACE est disponible pour  
caractériser l'extrémité 5' manquante d'un cDNA (Fromont-Racine et al.  
Nucleic Acid Res.21(7) :1683-1684 (1993)). En bref, un oligonucléotide  
d'ARN est ligaturé à l'extrémité 5' d'une population d'ARNm. Après  
30 retrotranscription en cDNA, un jeu d'amorces spécifiques respectivement de  
l'adaptateur ligaturé en 5' et d'une séquence située en 3' du gène d'intérêt  
est utilisé en PCR pour amplifier la portion 5' du cDNA recherché. Le  
fragment amplifié est ensuite utilisé pour reconstruire l'ADNc complet.

### EXEMPLE 3 : Analyse du profil d'expression génique pour la maladie de Tangier

La vérification de l'altération du niveau d'expression du gène ABC1 entraînant le phénotype cellulaire de Tangier peut-être déterminé par hydridation de ces séquences avec des sondes correspondant aux ARNm provenant de fibroblastes de sujets atteints ou non de la maladie, selon les méthodes décrites ci-dessous :

#### 10 1. Préparation des ARN totaux, des ARNm poly(A)<sup>+</sup> et de sondes de cDNA

Les ARN totaux sont obtenus à partir de cultures cellulaires des fibroblastes de sujets normaux ou bien atteints de la maladie de Tangier par la méthode à l'isothiocyanate de guanidine (Chomczynski & Sacchi, 1987). Les ARNm poly(A)<sup>+</sup> sont obtenus par chromatographie d'affinité sur colonnes d'oligo(dT)-cellulose (Sambrook et al., 1989) et les cDNA utilisés comme sondes sont obtenus par RT-PCR (DeRisi et al., 1997) avec des oligonucléotides marqués avec un produit fluorescent (Amersham Pharmacia Biotech ; CyDye™).

#### 20 2. Hydridation et détection des niveaux d'expressions

Les membranes de verre contenant les séquences présentées dans cette demande de brevet, correspondant au gène Tangier sont hybridées avec les sondes de cDNA, obtenues à partir des fibroblastes (Iyer et al., 1999). L'utilisation du système Amersham/molecular Dynamics (Avalanche Microscanner™) permet la quantification des expressions des produits de séquences sur le type cellulaire sain ou affecté.

### EXEMPLE 4 : Construction du vecteur d'expression contenant l'ADNc complet de ABC1 dans des cellules de mammifères

30

Le gène ABC1 peut être exprimé dans des cellules de mammifères. Un vecteur d'expression eucaryote typique contient un promoteur qui permet l'initiation de la transcription de l'ARNm, une séquence codant pour la protéine, et les signaux requis pour la terminaison de la transcription et pour la polyadénylation du transcrit. Il contient aussi des signaux supplémentaires

35

l'animal dans le but de produire des antisera polyclonaux de plus grande activité.

Dans la méthode préférée, les anticorps de la présente invention sont des anticorps monoclonaux. De tels anticorps monoclonaux peuvent être  
5 préparés en utilisant la technique d'hybridome. (Köhler et al, Nature 256 :495 (1975) ; Köhler et al, Eur. J. Immunol. 6 :511 (1976) ; Köhler et al, Eur. J. Immunol. 6:292 (1976) ; Hammeling et al., in : Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 51981). En général, de telles méthodes impliquent d'immuniser l'animal (préférentiellement une souris)  
10 avec un polypeptide ou, mieux encore, avec une cellule exprimant le polypeptide. Ces cellules peuvent être mises en culture dans un milieu de culture tissulaire adéquat. Cependant, il est préférable de cultiver les cellules dans un milieu Eagle (Earle modifié) supplémenté avec 10% de serum bovin foetal (inactivé à 56°C) et additionné d'environ 10 g/l d'acides aminés non  
15 essentiels, de 1000 U/ml de pénicilline et d'environ 100 µg/ml de streptomycine.

Les splénocytes de ces souris sont extraits et fusionnés avec une lignée cellulaire de myelome adéquate. Cependant, il est préférable d'utiliser la lignée cellulaire de myelome parentale (SP2O) disponible à l'ATCC. Après  
20 fusion, les cellules d'hybridomes résultantes sont sélectivement maintenues en milieu HAT puis clonées par dilution limite comme décrit par Wands et al. (Gastroentérolgy 80:225-232 (1981)). Les cellules d'hybridomes obtenues après une telle sélection sont testées afin d'identifier les clones sécrétant des anticorps capables de se fixer au polypeptide.

25 D'autre part, d'autres anticorps capables de se fixer au polypeptide peuvent être produits selon une procédure en 2 étapes utilisant des anticorps anti-idiotypique une telle méthode est fondée sur le fait que les anticorps sont eux-mêmes des antigènes et en conséquence il est possible d'obtenir un anticorps reconnaissant un autre anticorps. Selon cette  
30 méthode, les anticorps spécifiques de la protéine sont utilisés pour immuniser un animal, préférentiellement une souris. Les splénocytes de cet animal sont ensuite utilisés pour produire des cellules hybridomes, et ces dernières sont criblées pour identifier les clones qui produisent un anticorps dont la capacité à se fixer au complexe protéine-anticorps spécifique peut-  
35 être bloqué par le polypeptide. Ces anticorps peuvent être utilisés pour

immuniser un animal afin d'induire la formation en plus grande quantité d'anticorps spécifiques de la protéine.

Il serait apprécié que Fab et F(ab')<sub>2</sub> et les autres fragments des anticorps de la présente invention puissent être utilisés selon les méthodes décrites ici. De tels fragments sont typiquement produits par clivage protéolytique à l'aide d'enzymes telles que la Papaïne (pour produire les fragments Fab) ou la Pepsine (pour produire les fragments F(ab')<sub>2</sub>). Sinon, les fragments sécrétés reconnaissant la protéine peuvent être produits en appliquant la technologie de l'ADN recombinant ou de la chimie de synthèse.

10 Pour l'utilisation in vivo d'anticorps chez l'homme il serait préférable d'utiliser des anticorps monoclonaux chimériques "humanisés". De tels anticorps peuvent être produits en utilisant des constructions génétiques dérivés de cellules d'hybridomes produisant les anticorps monoclonaux décrits ci-dessus. Les méthodes pour produire les anticorps chimériques  
15 sont connus par l'homme de l'art. (Pour revue, voir : Morrison, Science 229 :1202 (1985) ; Oi et al., Biotechnique 4 :214 (1986) ; Cabilly et al., US patent n°4,816,567 ; Taniguchi et al., EP 171496 ; Morrison et al., EP 173494 ; Neuberger et al., WO 8601533 ; Robinson et al., WO 8702671 ; Boulianne et al ; Nature 312 :643 (1984) ; Neuberger et al., Nature 314 : 268  
20 (1985)).

#### **EXEMPLE 7 : Correction du phénotype cellulaire de la maladie de Tangier**

25

La maladie de Tangier est caractérisée par un catabolisme accéléré des particules lipoprotéiques de haute densité (HDL) et une accumulation de cholestérol dans les tissus. Notamment, les fibroblastes de peau des patients atteints de la maladie de Tangier ont une capacité réduite à éliminer  
30 leur contenu en cholestérol par le processus d'efflux de cholestérol assuré par l'apolipoprotéine A-I (apoA-I), protéine majeure des HDL (Francis et al., 1995). Cette caractéristique correspondant à une perte de fonction est aussi retrouvée dans d'autres cellules fibroblastiques de patients atteints de déficit familial en HDL (Marcil et al., 1999).

35



La correction du phénotype des fibroblastes de Tangier peut être assurée par la transfection de l'ADNc complet de ABC1 selon l'invention, dans lesdites cellules. L'ADNc est inséré dans un vecteur d'expression qui est ensuite transfecté selon les méthodes décrites ci-dessous :

5

1. Préparation des cultures fibroblastiques de sujets normaux et de sujets atteints de la maladie de Tangier

Les fibroblastes primaires de peau humaine sont obtenus par la mise en culture de biopsie de peau provenant de l'avant bras. Ces biopsies sont effectuées sur des patients atteints de la maladie de Tangier ayant les particularités cliniques et biochimiques des " homozygotes ", c'est à dire des amygdales oranges, des concentrations plasmatiques d'apoA-I et de cholestérol-HDL inférieur au 5<sup>ème</sup> percentile. Les lignées de fibroblastes normaux sont obtenus chez l'American Type Culture Collection (Rockville, MD). Les fibroblastes sont cultivés dans un milieu EMMEM (Eagle-modified minimum essential medium ; GIBCO) complété par 10% de sérum de veau foetal, de la glutamine à 2 mM, 100 UI/ml de pénicilline et 100 µg/ml de streptomycine (milieu désigné par EMMEM10). En vue de la réalisation de l'étude de l'efflux de cholestérol, ces cellules sont pré-chargées en cholestérol par incubation de 24 heures avec 50µg/ml de cholestérol dans le milieu décrit ci-dessus sans sérum de veau mais contenant 2 mg/ml d'albumine bovine (BSA, fraction V).

2. Etude de l'efflux de cholestérol

Les fibroblastes pré-chargés en cholestérol à confluence sur des plaques à 24 puits sont incubés dans le milieu EMMEM10 et 1 µCi/ml de 1,2-<sup>3</sup>H-cholestérol (50 Ci/mmol ; Dupont ; Wilmington, DE) durant 48 heures. Environ 100 000 coups par minute sont obtenus par puits ou 1000 coups par minute et par µg de protéine cellulaire. Les cellules sont lavées trois fois avec du milieu EMMEM/BSA, et incubées avec ce milieu durant 24 heures avant de transfecter le gène d'intérêt et de démarrer l'efflux par ajout de 10 µg/ml de protéoliposome contenant l'apoA-I en milieu EMMEM/BSA. Ces protéoliposomes sont préparés par sonication de phosphatidylcholine et d'apoA-I humaine purifiée (Jonas, 1986). La transfection cellulaire s'effectue par la technique de précipitation de phosphate de calcium (Sambrook et al.,

1989). Après la période d'efflux, en général 20 heures, le milieu est collecté, centrifugé (1000 g, 5 min), et la radioactivité déterminée par comptage en scintillation liquide. La radioactivité résiduelle dans les cellules est aussi déterminée sur la nuit après extraction des lipides dans l'isopropanol. Le  
5 pourcentage d'efflux est calculé en divisant la radioactivité mesurée dans le surnageant par la somme des radioactivités mesurées, dans le surnageant et l'extrait cellulaire. Un contrôle interne est réalisé par transfection d'un gène marqueur et l'incubation sur 24 heures avec un milieu EMMEM/BSA sans protéoliposome contenant l'apoA-I. L'efflux de cholestérol cellulaire à partir  
10 de fibroblastes normaux et transfectés par un gène témoin correspondent à  $6 \pm 2\%$  alors que celui obtenu à partir de fibroblastes atteints de la maladie de Tangier et transfectés par ce gène témoin est inférieur à 1%. En revanche la transfection des fibroblastes atteints de la maladie de Tangier par un plasmide contenant l'ADNc complet ou encore l'ADN génomique de ABC1  
15 selon l'invention permettrait de restaurer la capacité de ces cellules à éliminer leur excès de cholestérol à un niveau correspondant à celui de fibroblastes normaux.

**EXEMPLE 8 : Isolement et caractérisation de fragments génomiques du**  
20 **gène ABC1 humain.**

Un fragment d'environ 3 kb de l'ADNc de ABC1 humain a été obtenu à partir du clone d'ADNc désigné "pf10" contenant le premier domaine de fixation de l'ATP de ABC1, ce clone d'ADNc étant décrit dans l'article de  
25 Luciani et al. (1994).

Ce fragment d'ADNc obtenu par digestion du clone pf10 à l'aide de l'endonucléase de restriction EcoRI, a été isolé sur un gel d'agarose après électrophorèse, puis marqué avec la digoxigénine selon les instructions du constructeur (kit commercialisé par Boehringer Mannheim, référence 1 585  
30 614)

Le fragment d'ADNc marqué a été utilisé afin de cribler la banque de cosmides LLNL (Lawrence Livermore National Labs) du chromosome 9, immobilisés sur un filtre de Nylon<sup>TM</sup>.

Six clones positifs ont été identifiés. Pour ces six cosmides, la sonde  
35 a hybridé avec des colonies uniques.

Un clone représentatif a été isolé à partir de chacune de ces colonies.

Les clones LLNLC 131J087 Q2 (désignés ici cos3a) et LLNLC 131O1165 Q2 (désignés ici cos6f) ont été analysés plus en détail.

Le clone cos3a a été sous-cloné sous la forme d'un fragment EcoRI  
5 dans le vecteur Gen3zf(-) et séquencé aux deux extrémités en utilisant la technologie Big Dye Terminator sur un séquenceur de type ABI377 (Applied Biosystems, Perkin Elmer).

Les clones contenant des inserts distincts (déterminés après  
séquençage des extrémités des différents inserts ou encore par  
10 détermination de la taille de ceux-ci) qui étaient de longueur trop grande pour être complètement séquencés à l'aide des amorces hybridant avec les séquences du vecteur, ont été analysés plus avant par la technique d'insertion de transposon puis de séquençage à l'aide d'amorces spécifiques du transposon (système " GPS " commercialisé par la Société New England  
15 Biolabs).

De cette manière, des séquences génomiques correspondant au gène ABC1 humain ont été isolées et caractérisées. Ces séquences ont été comparées avec des séquences humaines et de souris référencées dans les bases de données permettant de déterminer les jonction intron-exon..

20

#### **EXEMPLE 9 : Détermination de polymorphismes/mutations dans le gène ABC1.**

La détection de polymorphismes et ou de mutations dans les séquences des transcrits ou dans la séquence génomique du gène ABC1 peut être réalisée  
25 selon différents protocoles. La méthode de choix est le séquençage direct.

Pour les patients dont on peut obtenir une préparation d'ARNm la méthode de choix consiste à préparer les cDNAs et les séquencer directement, Pour les patients dont on ne dispose que de l'ADN, et dans le cas d'un transcrit où  
30 la structure du gène correspondant est inconnue ou partiellement connue il est nécessaire de déterminer précisément sa structure intron-exon ainsi que la séquence génomique du gène correspondant. Il s'agit donc dans un premier temps d'isoler le ou les clones de cosmides ou de BAC d'ADN génomique correspondant au transcrit étudié selon la méthode décrite dans  
35 l'exemple 8, de séquencer l'insert du ou des clones correspondants et de

déterminer la structure intron-exon en comparant la séquence de l'ADNc à celle de l'ADN génomique obtenu.

La technique de détection de mutations par séquençage direct consiste à comparer les séquences génomiques du gène ABC1 obtenues à partir des  
5 homozygotes pour la maladie ou d'au moins 8 individus ( 4 individus affectés par la pathologie étudiée et 4 individus non affectés). Les divergences de séquence constituent des polymorphismes. Tous ceux modifiant la séquence en acides aminés de la protéine sauvage peuvent être des mutations susceptibles d'affecter la fonction de ladite protéine qu'il est intéressant de  
10 considérer plus particulièrement pour l'étude de co-ségrégation de la mutation et de la maladie (corrélation génotype-phénotype) dans le pédigrée ou encore dans les études d'association cas/témoin pour l'analyse des cas sporadiques.

15 **EXEMPLE 10 : Identification d'un gène causal d'une maladie liée à un déficit du transport inverse du cholestérol par la mutation causale ou une différence transcriptionnelle**

Parmi les mutations identifiées selon la méthode décrite dans l'Exemple 9, toutes celles associées au phénotype malade sont susceptibles d'être  
20 causales. La validation de ces résultats est faite en séquençant le gène chez tous les individus atteints et leurs apparentés (dont l'ADN est disponible).

D'autre part, la réalisation de Northern blot ou RT-PCR, selon la méthode décrite dans l'Exemple 1, à partir d'ARN spécifique d'individus atteints et non-atteints permet de détecter des variations notables du niveau  
25 d'expression du gène étudié, en particulier une absence de transcription du gène.

**EXEMPLE 11 : Identification d'une délétion d'un nucléotide dans l'exon 13 du gène ABC1 chez des patients atteints de la maladie de TANGIER**

30 L'analyse de mutations dans le gène ABC1 a été réalisée sur de l'ADN génomique de plusieurs individus appartenant à une famille dont plusieurs membres sont atteints de la maladie de Tangier avec des désordres coronariens précoces.

Une délétion d'un nucléotide a été identifiée dans l'exon 13 (DG  
35 1764 : Leu548Leu ;575 End). Cette délétion introduit un codon stop en

position 575 qui permet de prédire une troncation de la protéine ABC1 codée par le gène ABC1 muté, cette troncation conduisant à la synthèse d'un polypeptide délété d'une grande partie de la séquence normale d'acides aminés, et en particulier des deux cassettes de fixation à l'ATP.

- 5 Une corrélation parfaite entre l'observation des symptômes de la maladie et la présence de cette délétion d'un nucléotide a été retrouvée dans l'ensemble de la famille (Figure 1).

**EXEMPLE 12 : Identification d'une insertion d'un segment de**  
10 **nucléotides dans l'exon 12 du gène ABC1.**

Dans une autre famille dont plusieurs membres sont atteints de la maladie de Tangier, on a observé une insertion de 110 paires de bases ayant la structure d'une séquence nucléotidique répétée de type Alu-sq, accompagnée d'une délétion de 14 paires de bases dans l'exon 12 (Figure  
15 2). Cette mutation par insertion/délétion permet de prédire une délétion de 65 acides aminés (DERKFW) ainsi qu'une insertion en phase de 38 acides aminés (EYSGVTSAHCNLCLLSSSDSRASASQVAGITAPATTPG).

Cette mutation ne permet pas la synthèse d'un polypeptide transporteur ABC1 normal. Il peut donc être conclu que la maladie de  
20 Tangier, chez les individus de cette famille, est provoquée par un défaut dans le gène ABC1.

**EXEMPLE 13. : Identification de polymorphismes bialléliques dans le**  
gène ABC1

25 Des amorces pour l'amplification de l'ADN des patients ont été conçues à partir des séquences non répétitives de l'ADN intronique du gène ABC1, de telle manière à ce qu'une amplification des jonctions intron-exon ainsi que des bases essentielles pour la formation de la structure secondaire durant l'étape d'épissage de l'ARN soit incluses dans les fragments  
30 amplifiés.

Les différents couples d'amorces spécifiquement mis au point sont présentés dans le tableau V.

Les résultats trouvés sur l'ADN d'une famille contenant des cas de maladie de Tangier sans complication coronariennes sont montrés dans le tableau  
35 IV.

L'ADN génomique des patients a été amplifié à l'aide des amorces décrites ci-dessus en utilisant le kit Qiagen's Star Taq ou encore le kit Supertaq, en utilisant les conditions d'hybridation et des conditions de cycle d'amplification préconisées par le constructeur.

5 Les produits PCR amplifiés ont été purifiés en utilisant un kit commercialisé par la Société Qiagen, puis séquencés par la méthode Big Dye Terminator sur un séquenceur ABI377 (Applied Biosystems, Perkin Elmer).

10 **EXEMPLE 14 : Identification d'une région de 1 cM sur le locus 9q31 associée à la maladie de Tangier**

Une première analyse de liaison ("linkage") a été décrite dans l'article de Rust et al. (1998).

15 Cet article a présenté une analyse de liaison sur trois familles de patients atteints de la maladie de Tangier et défini un intervalle candidat de 8 cM en 9q31.

Le demandeur a réalisé une étude de liaison en incluant quatre familles supplémentaires ainsi que des marqueurs additionnels référencés  
20 dans des bases de données publiques afin d'affiner la région candidate à environ 1 cM, en référence à la carte génétique publiée par le Généthron (Dib et al., 1996)

Les résultats d'analyse de liaison présentés ci-après ont permis au demandeur d'exclure de la région candidate les segments génomiques  
25 respectivement proximaux (centromérique) et distaux (télomérique) aux marqueurs D9S271 et D9S1866.

La région candidate est donc localisée entre ces deux marqueurs exclus.

Une information importante ayant permis d'affiner la région candidate  
30 a été obtenue à partir de la portion E1 du pedigree décrit dans l'article de Rust et al. (1998). En effet il a été montré sur le chromosome maternel, à l'origine de la partie E121m, que l'événement de recombinaison (crossing-over) déjà décrit dans la figure 2 de cet article (entre les marqueurs D9S277 et D9S53), devait en fait être localisé de façon télomérique par rapport au

marqueur D9S271. La borne centromérique de l'intervalle candidat étant située entre D9S271 et D9S277.

Dans deux des nouvelles familles, respectivement la famille " S1 " et la famille " NU ", des événements de recombinaison supplémentaires ont été observés. Ces événements de recombinaison ont permis de déplacer la borne télomérique de l'intervalle candidat, du marqueur D9S1677 (décrit dans l'article de Rust et al 1998) au marqueur D9S1866.

Le premier pedigree, " S1 ", a été étendu. Les individus atteints présentent un génotype homozygote pour tous les marqueurs de la région de 8cM ainsi que pour des marqueurs plus lointains, localisés de part et d'autre de cette région. L'un des cousins de l'individu S1, apparenté à S1 par les deux parents (double consanguinité) possède quatre enfants. Deux de ces enfants présentent sur leur chromosome d'origine paternel une conservation d'une grande partie de l'haplotype malade (dans la région défini de 8 cM). Ces deux enfants présentent également la caractéristique typique des parents hétérozygotes de la famille atteinte de la maladie de Tangier, à savoir un niveau de HDL qui est de moitié le niveau observé chez des patients non affectés par la maladie.

Cependant, le caractère homozygote pour les marqueurs n'est plus observé dans la région chromosomique à partir du marqueur D9S1866 (qui est hétérozygote chez ces individus), ce qui a permis de définir D9S1866 comme borne télomérique de la région candidate.

La même borne télomérique a été observée dans la famille " Nu ", dans laquelle l'un des quatre enfants issus de parents cousins au premier degré, était un patient affecté de la maladie de Tangier homozygote.

Une homozygotie pour les marqueurs sur l'ensemble de la région candidate a été observée chez ce patient.

L'un de ses frères, hétérozygote (au niveau phénotypique), présente un événement de recombinaison (crossing-over) sur l'un des deux chromosomes, de telle manière que ce frère est homozygote (au niveau génétique) pour tous les marqueurs télomériques incluant le marqueur D9S1866, mais hétérozygote (au niveau génétique) pour les marqueurs localisés dans la région vers le centromère.

Comme ce patient ne présentait pas un phénotype de patient homozygote, la région télomérique incluant le marqueur D9S1866 a pu être exclue.

**5 EXEMPLE 15: Isolement et caractérisation du gène ABC1 humain.**

Un fragment d'environ 3 kb de l'ADNc de ABC1 humain a été obtenu à partir du clone d'ADNc désigné "pf10" contenant le premier domaine de fixation de l'ATP de ABC1, ce clone d'ADNc étant décrit dans l'article de  
10 Luciani et al. (1994).

Ce fragment d'ADNc obtenu par digestion du clone pf10 à l'aide de l'endonucléase de restriction EcoRI, a été isolé sur un gel d'agarose après électrophorèse, puis marqué avec la digoxigénine selon les instructions du constructeur (kit commercialisé par Boehringer Mannheim)

15 Le fragment d'ADNc marqué a été utilisé afin de cribler la banque de cosmides LLNL du chromosome 9, immobilisés sur un filtre de Nylon<sup>TM</sup>.

Six clones positifs ont été identifiés. Pour ces six cosmides, la sonde a hybridé avec des colonies uniques.

Un clone représentatif a été isolé à partir de chacune de ces colonies.

20 Les clones LLNLC 131J087 Q2 (désignés ici cos3a) et LLNLC 131O1165 Q2 (désignés ici cos6f) ont été analysés plus en détail.

Le clone cos3a a été sous-cloné sous la forme d'un fragment EcoRI dans le vecteur Gen3zf(-) et séquencé aux deux extrémités en utilisant la technologie Big Dye Terminator sur un séquenceur de type ABI377.

25 Les clones contenant des inserts distincts (déterminés après séquençage des extrémités des différents inserts ou encore par détermination de la taille de ceux-ci) qui étaient de longueur trop grande pour être complètement séquencés à l'aide des amorces hybridant avec les séquences du vecteur, ont été analysés plus avant par la technique  
30 d'insertion de transposon puis de séquençage à l'aide d'amorces spécifiques du transposon (système "GPS" commercialisé par la Société New England Biolabs).

De cette manière, des séquences génomiques correspondant au gène ABC1 humain ont été isolées et caractérisées. Ces séquences ont été



comparées avec des séquences humaines et de souris référencées dans les bases de données.

Les séquences des jonctions intron-exon ont été déterminées.

Des amorces pour l'amplification de l'ADN des patients ont été  
5 conçues à partir des séquences non répétitives de l'ADN intronique du gène ABC1, de telle manière à ce qu'une amplification des jonctions intron-exon ainsi que des bases essentielles pour la formation de la structure en lariat durant l'étape d'épissage soit incluse dans les fragments amplifiés.

L'ADN génomique des patients a été amplifié à l'aide des amorces  
10 décrites ci-dessus en utilisant le kit Qiagen's Star Taq ou encore le kit Supertaq, en utilisant les conditions d'hybridation et des conditions de cycle d'amplification préconisées par le constructeur.

Les produits PCR amplifiés ont été purifiés en utilisant un kit commercialisé par la Société Qiagen, puis séquencés par la méthode Big  
15 Dye Terminator sur un séquenceur ABI377.

#### **EXEMPLE 16: Construction de vecteurs recombinants contenant un polynucléotide codant pour la protéine ABC1**

##### **I. Synthèse de du gène ABC1 humain.**

20

De l'ARN total (500 ng) isolé de tissu de placenta humain (Clontech, Palo Alto, CA, USA) a été utilisé comme source pour la synthèse de l'ADNc du gène ABC1 humain, en mettant en oeuvre le système " Superscript one step RT-PCR (Life Technologies, Gaithersburg, MD, USA ) et les amorces  
25 oligonucléotidiques spécifiques de ABC1 (0,25 µM) suivante:

- amorce aller : 5'-CTACCCACCCTATGAACAAC-3' (nt 75-94 de l'ADNc ABC1);

- amorce retour: 5'-TCCACCCCGTATGAACAGGG-3' (nt 6731-6751 de l'ADNc de ABC1).

30 Ces amorces oligonucléotidiques ont été synthétisées par la méthode au phosphoramidite sur un appareil synthétiseur d'ADN de type ABI 394 (Applied Biosystems, Foster City, CA, USA).

Les sites reconnus par l'enzyme de restriction NotI ont été incorporés dans l'ADNc amplifié de 6676 pb par une nouvelle étape d'amplification en  
35 utilisant 50 ng de l'ADNc de ABC1 humain comme matrice, et 0,25 µM des

amorce oligonucléotidiques décrites ci-dessus contenant, à leur extrémité 5', le site reconnu par l'enzyme de restriction NotI, en présence de 200µM de chacun desdits didéoxynucléotides dATP, dCTP, dTTP et dGTP ainsi que l'ADN polymérase de *Pyrococcus furiosus* (Stratagene, Inc. La Jolla, CA, 5 USA).

La réaction PCR a été réalisée pendant 30 cycles comprenant chacun une étape de dénaturation à 95°C pendant une minute, une étape de renaturation à 50°C pendant une minute et une étape d'extension à 72°C pendant deux minutes, dans un appareil thermocycleur pour PCR (Cetus 10 Perkin Elmer Norwalk, CT, USA).

## **II. Clonage de l'ADNc du gène ABC1 humain dans un vecteur d'expression:**

L'insert de 6676 pb de l'ADNc de ABC1 humain a été cloné au site de restriction NotI du vecteur d'expression pCMV contenant un promoteur précoce du cytomégalovirus et une séquence activatrice ("Enhancer") ainsi que le signal de polyadénylation de SV40 (Beg et al., 1990; Applebaum-Boden, 1996), afin de réaliser le vecteur d'expression désigné pABC1.

La séquence de l'ADNc cloné a été confirmée par un séquençage sur les deux brins en utilisant la trousse de réaction "ABI Prism Big Dye Terminator Cycle Sequencing ready" (commercialisée par Applied Biosystems, Foster City, CA, USA) dans un séquenceur capillaire de type ABI 310 (Applied Biosystems, Foster City, CA, USA).

## **III. Construction d'un vecteur adénoviral recombinant contenant l'ADNc du gène ABC1 humain,**

### **A- Modification du vecteur d'expression pCMV- $\beta$ .**

L'ADNc de la  $\beta$ -galactosidase du vecteur d'expression pCMV- $\beta$  (Clontech, Palo Alto, CA, USA, Gene Bank Accession n°U02451) a été délété par digestion avec l'endonucléase de restriction NotI et remplacé par un polysite de clonage contenant, de l'extrémité 5' vers l'extrémité 3', les sites suivants:

NotI, AscI, RsrII, AvrII, Swal, et NotI (séquence du polysite de clonage:

5'-CGGCCGCGGCGCGCCCGACCGCCTAGGATTAAATCGCGGCCCGCG-3'

ce polysite de clonage ayant été cloné au niveau du site NotI.

Le fragment d'ADN compris entre les sites EcoRI et SmaI du vecteur d'expression pCMV modifié a été isolé et cloné dans le site XbaI modifié du vecteur navette pXCXII (McKinnon et al., 1982; McGrory et al., 1988).

### **B-modification du vecteur navette pXCXII.**

Un polysite de clonage comprenant, de l'extrémité 5' vers l'extrémité 3' les sites de restriction XbaI, EcoRI, SfiI, PmeI, NheI, SrfI, PacI, Sall et XbaI (de séquence:

5' - GCTCTAGAATTCGGCCTCCGTGGCCGTTTAAACGCTAGCGCCCGGGCTTAATTAA  
GTCGACTCTAGAGC-3')

a été insérée au niveau du site XbaI (nucléotide en position 3329) du vecteur pXCXII (McKinnon et al., 1982; McGrory et al., 1988).

Le fragment d'ADN EcoRI-Sall isolé du vecteur pCMV- $\beta$  modifié contenant le promoteur/enhancer de CMV, les sites donneurs et accepteurs d'épissage de FV40 et le signal de polyadénylation de FV40 a ensuite été cloné dans le site EcoRI-Sall du vecteur navette pXCX modifié, désigné pCMV-11.

#### **C- Préparation du vecteur navette pAD12-ABC1.**

L'ADNc ABC1 humain est obtenu par une réaction de RT-PCR, comme décrit ci-dessus, et cloné au niveau du site NotI dans le vecteur pCMV-12, résultant dans l'obtention du vecteur pCMV-ABC1.

L'ADNc ABC1 contenu dans le vecteur pCMV-ABC1 est constitué d'un fragment d'ADN de 6676 pb comprenant la séquence allant du nucléotide en position 75 au nucléotide en position 6751 de l'ADNc ABC1 humain.

#### **D. Construction de l'adénovirus recombinant ABC1.**

L'adénovirus recombinant ABC1-rldV contenant l'ADNc ABC1 humain a été construit selon la technique décrite par McGrory et al. (1988).

Brièvement, le vecteur pAD12-ABC1 a été cotransfecté avec le vecteur tGM17 selon la technique de CHEN et OKAYAMA (1987).

De même, le vecteur pAD12-Luciférase a été construit et co-transfecté avec le vecteur pJM17.

Les adénovirus recombinants ont été identifiés par amplification PCR et soumis à deux cycles de purification avant une amplification à grande échelle dans la lignée de cellules embryonnaires de rein humain HEK 293 (American Type Culture Collection, Rockville, MD, USA).

5 Les cellules infectées ont été collectées 48 à 72 heures après leur infection par les vecteurs adénoviraux et soumis à cinq cycles de lyse par congélation et décongélation.

Les lysats bruts ont été extraits à l'aide de Fréon (Halocarbène 113, Matheson Product, Scaucus, N.J. USA), sédimentés deux fois dans le  
10 chlorure de césium supplémenté d'albumine murine à 0,2% (Sigma Chemical Co., Saint-Louis, MO, USA) et dialysés extensivement contre du tampon composé de 150 mM NaCl, 10 mM Hepes (pH 7,4), 5 mM KCl, 1 mM MgCl<sub>2</sub>, et 1 mM CaCl<sub>2</sub>.

Les adénovirus recombinants ont été conservés à -70°C et titrés avant  
15 leur administration à des animaux ou leur incubation avec des cellules en culture.

L'absence d'adénovirus contaminant de type sauvage a été confirmée par criblage à l'aide d'amplification PCR en utilisant des amorces oligonucléotidiques localisées sur la partie structurale de la région délétée.

20

#### **IV Validation de l'expression de l'ADNc ABC1 humain.**

Des anticorps polyclonaux spécifiques du polypeptide ABC1 humain ont été préparés chez des lapins et des poussins par injection du peptide  
25 synthétique "LHKNQTVVDVAVLTSFLQDEKVKESYV", dérivé de la protéine ABC1. Ces anticorps polyclonaux sont utilisés pour détecter et/ou quantifier l'expression du gène ABC1 humain dans des cellules et des modèles animaux par immunoempreinte et/ou immunodétection.

L'activité biologique de ABC1 peut être suivie en quantifiant les flux de  
30 cholestérol induits par apoA-I à partir de cellules transfectées par le vecteur pCMV-ABC1 qui ont été chargées en cholestérol (Remaley et al., 1997).

#### **V. Expression in vitro de l'ADNc abc1 humain dans les cellules.**

Des cellules de la lignée HEK293 et de la lignée COS-7 (American Tissue Culture Collection, Bethesda, MD, USA), ainsi que des fibroblastes en culture primaire dérivés de patients Tangier ou de patients atteints d'hypo-alphalipoprotéïnémie sont transfectés avec le vecteur d'expression pCMV-ABC1 (5-25µg) en utilisant de la Lipofectamine (BRL, Gaithersburg, MD, USA) ou par co-précipitation à l'aide de chlorure de calcium (Chen et al., 1987).

Ces cellules peuvent être également infectées avec le vecteur pABC1-AdV (Indice d'infection, MOI=10).

L'expression de ABC1 humain peut être suivie par immunoempreinte ainsi que par quantification de l'efflux de cholestérol induit par apoA-1 à partir de cellules transfectées et/ou infectées.

La complémentation du défaut génétique dont sont atteints les patients Tangier et les patients hypo-alphalipoprotéïnémiques à partir de fibroblastes de ces patients, peut être confirmée par la détection de l'expression du gène ABC1 normal, ce qui permet d'établir l'importance fonctionnelle de ce récepteur.

#### **VI. : Expression in vivo du gène ABC1 dans différents modèles animal.**

Un volume approprié (100 à 300 µl) d'un milieu contenant l'adénovirus recombinant purifié (pABC1-AdV ou pLucif-AdV) contenant de  $10^8$  à  $10^9$  unités formant des plages de lyse (PFUs) sont infusées dans la veine Saphène de souris (C57BL/6, à la fois souris témoins et modèles de souris transgéniques ou knock-out) au jour 0 de l'expérience.

L'évaluation du rôle physiologique de la protéine ABC1 dans le métabolisme des lipoprotéines est réalisée par détermination de la quantité totale de cholestérol, de triglycérides, de phospholipides et de cholestérol libre (Sigma et Wako Chemicals, Richmond, VA, USA), de cholestérol-HDL (CIBA-Corning, Oberlin, OH, USA) et des apolipoprotéines A-I, A-II, E et B de souris (Foger et al., 1997), avant (jour zéro) et après (jours 2, 4, 7, 10, 14) l'administration de l'adénovirus.

Des études cinétiques à l'aide de produits marqués radioactivement tels qu'apoA-I-HDL, CE-HDL ainsi que apoB-LDL et CE-LDL sont réalisées au jour 5 après l'administration des vecteurs rLucif-AdV et rABC1-AdV afin

d'évaluer l'effet de l'expression de ABC1 sur le métabolisme des HDLs et des LDLs ainsi que sur la libération du cholestérol vers le foie.

L'effet de l'expression de ABC1 sur le développement de l'athérosclérose peut être évalué en quantifiant la surface moyenne de lésion aortique dans des souris apoE après administration du vecteur rABC1-Adv.

De plus, des souris transgéniques et des lapins surexprimant le gène ABC1 peuvent être produits, conformément à l'enseignement de Waisman (1995) et Hoeg (1996) en utilisant des constructions contenant l'ADNc de ABC1 humain sous le contrôle de promoteurs endogènes tels que ABC1, CMV ou apoE.

L'évaluation de l'effet à long terme de l'expression de ABC1 sur la cinétique des lipides, lipoprotéines, apolipoprotéines plasmatiques et sur l'athérosclérose pourra être réalisée comme décrit ci-dessus.

**EXEMPLE 17 : Utilisation de vésicules pour le criblage de molécules agonistes et antagonistes de la protéine ABC1.**

La base de cette essai est la reconstitution de membranes ayant incorporé la protéine ABC1 et contenant des substrats comme le cholestérol ou des phospholipides. La protéine ABC1 peut alors être activée ou sa fonction réprimée par l'ajout de molécules d'intérêt. La sortie des substrats par le canal formé par la protéine ABC1 est alors détectée.

*a) Reconstitution d'une membrane contenant la protéine ABC1 et un substrat lipidique marqué.*

Différentes stratégies peuvent être employées pour fabriquer ces membranes, des méthodes utilisant des solvant organiques, des moyens mécaniques comme la sonication, la " French press ", ou bien par des cycles de congélation- décongélation, ou encore utilisant des détergents (les cholates, le Chaps, Chapso) (référence : Rigaud et al. Biochimica et Biophysica Acta 1231 (1995) 223-246).

Plus particulièrement, un substrat lipidique comme des phospholipides, du cholestérol ou ester de cholestérol, radioactif du type 3H-cholestérol, 125-I-cholestérol, 3H-phosphatidylcholine ou bien fluorescent avec du NBD ou du pyrene (Molecular Probes ; <http://www.probes.com>) et de la phosphatidyle-

choline d'œuf (1 mM) sont séchés sur le culot d'un flacon en verre. Dans ce flacon sont mélangés à la fois du cholate de sodium et la protéine ABC1 au ratio mole à mole de 0,3. L'ensemble est mélangé au vortex pendant 5 minutes puis incubé à 25°C durant 30 minutes puis dialysé contre un tampon salin. Le protéoliposome réalisé selon ce protocole est suivi par turbidimétrie pour vérifier sa bonne fabrication.

*b) Capture du protéoliposome sur une surface solide :*

Cette étape peut être réalisée en incorporant des protéines de fixation du type intégrine. Dans ce protocole, on utilise une capture par les anticorps dirigés contre la protéine ABC1 et préalablement adsorbée sur une plaque 96 ou 384 puits.

Une solution contenant ces anticorps à la concentration de 100 µg/l sont absorbés sur ces plaques multipuits par incubation sur la nuit à 4°C. Après lavage, la plaque est ensuite saturée par de l'albumine bovine à 1 mg/ml incubé durant 2 heures à 37°C. L'ensemble est encore lavé et incubé avec les protéoliposomes contenant ABC1 durant 2 heures à 37°C.

*c) Liaison aux molécules d'intérêt*

Cette étape est réalisée par incubation de produits durant 1 heure à 37°C

*d) Détermination de l'activation ou inhibition de la protéine ABC1*

Si le substrat est fluorescent, la fluorescence du surnageant nous révèle l'activité de produit à induire un transport de lipide vers l'extérieur du protéoliposome. Ou encore, l'utilisation de système Confocal nous informe sur les quantités de substrat à l'intérieur et l'extérieur du protéoliposome. Si le substrat est radioactif, l'utilisation de plaques du type CytoStar ayant un fond avec du liquide de scintillation permet de révéler le substrat encore séquestré dans le protéoliposome.

**EXEMPLE 18 : Utilisation de transport d'anion pour le criblage de molécules agonistes et antagonistes de la protéine ABC1).**

Le principe de cet essai réside dans la propriété qu'a la protéine ABC1 à transporter les anions lors de son activation.

a) Les cellules macrophagiques de la lignées THP-1, cellules humaines monocytic leukemia, sont un modèle de macrophages différenciés. Les cellules sont cultivées dans un milieu RPMI 1640 supplémenté par 10% de sérum de veau fœtal dans des plaques 48-multipuits à la densité de 2 105



cellules par puits. Les cellules fibroblastiques de patients atteints de la maladie de Tangier peuvent être utilisées comme témoin négatif parce que leur protéine ABC1 n'est pas fonctionnelle. Un autre témoin négatif peut être obtenu par l'ajout d'anticorps anti-ABC1.

5 b) L'utilisation de cellules défectives en transport anioniques ou bien des cellules traitées avec des inhibiteurs de canaux anioniques (type Verapamil, un inhibiteur de P-glycoprotéine ou le tetraethylammonium, un inhibiteur de canal potassique) peuvent être aussi utilisés.

c) Pour l'essai proprement dit, les cellules sont ensuite lavées par un milieu  
10 Earles's modified salt solution (ESS) préchargées avec 1ml de KI à 1  $\mu\text{mol/L}$  (0,1  $\mu\text{Ci/ml}$  de NaI125) dans ce milieu ESS pour 30 minutes à 37°C. Les produits sont alors ajoutés dans le milieu extracellulaire. Les cellules sont ensuite lavées par le milieu ESS.

d) La quantité d'iode dans le milieu est détectée toute les minutes durant  
15 11 minutes. Les deux premiers points correspondent à l'efflux basal. A la fin de l'incubation, le milieu est repris et la quantité d'iode restant dans les cellules est compté suite à la lyse des cellules dans du NaOH 1 molaire.

e) La quantité totale de radioactivité au temps zéro est égale à la somme de la radioactivité trouvée dans le surnageant et résiduelle dans les cellules.  
20 Les courbes d'efflux sont construites en marquant le pourcentage de radioactivité libéré dans le milieu en fonction du temps.

**EXEMPLE 19 : Utilisation de macrophages THP-1 exprimant l'IL-1bêta  
25 pour le criblage de molécules agonistes et antagonistes de la protéine ABC1).**

Le principe de ce test est que toutes substances modulant l'activité de la protéine ABC1 a des répercussions sur la synthèse d'IL-1beta.

30

a) Les cellules macrophagiques de la lignée THP-1, cellules humaines monocytic leukemia, sont un modèle de macrophages différenciés. Les cellules sont cultivées dans un milieu RPMI 1640 supplémenté par 10% de sérum de veau foetal dans des plaques multipuits à la densité de 2  $10^5$   
35 cellules par puits.

- b) Pour l'essai proprement dit, les cellules sont ensuite lavées et placés dans un milieu RPMI 1640 contenant 1 mg/ml d'albumine humaine purifiée fraction IV.
- c) Les produits sont ajoutés dans le milieu extracellulaire. Simultanément,  
5 les cellules sont alors activées par ajout de lipopolysaccharide (LPS) durant 3 heures à 1 µg/ml suivit par une incubation de 30 minutes en présence de ATP à 5 mmol/L.
- d) Les concentrations en IL-1beta et en contrôle l'IL-1alpha, le tumor necrosis factor alpha (TNFalpha) et l'IL-6 sont déterminées par des kits  
10 ELISA selon les instructions des fabricants (R&D Sytem ; IL-1bêta humain Chemiluminescent ELISA référence QLB00). Les variations en ARNm de l'IL-1bêta qui n'est pas censé être affecté sont évaluées par la technique de Northern blot avec la sonde correspondante.

## REFERENCES

- Altschul S.F. et al, J. Mol. Biol. 1990 215 : 403-410
- Altschul S.F. et al, Nucleic Acids Res. 1997 25 : 3389-3402
- Applebaum-Boden, JCI 97, 1996
- 5 Ausubel et al., 1989, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y
- Beard et al., Virology 75 (1990) 81
- Beaucage et al., *Tetrahedron Lett* 1981, **22**: 1859-1862
- Becq et al. Journal of Biological Chemistry vol 272, n°5 pages 2695-2699,  
10 1997
- Beg et al PNAS 87 p3473 1990
- Bender et al., J. Virol. 61 (1987) 1639
- Bernstein et al. Genet. Eng. 7 (1985) 235
- Boulianne et al ; Nature 312 :643 (1984)
- 15 Breakfield et al., New Biologist 3 (1991) 203
- Brown EL, Belagaje R, Ryan MJ, Khorana HG, *Methods Enzymol* 1979;**68**:109-151
- Cabilly et al., US patent n°4,816,567
- Chen and Kwok Nucleic Acids Research 25:347-353 1997
- 20 Chen et al. Proc. Natl. Acad. Sci. USA 94/20 10756-10761, 1997
- Chen et al., 1987, Mol. Cell. Biol., **7** : 2745-2752.
- Chomczynski, P. and Sacchi, 1987, Anal. Biochem., **162**, 156-159.
- DeRisi J. et al., 1997, Science, **278**, 680-686.
- Dib C. et al., 1996, Nature, **380** : 152-154.
- 25 Felgner et al., PNAS 84 (1987) 7413
- Flotte et al., 1992, *Am. J. Respir. Cell Mol. Biol.*, **7** : 349-356.
- Forsell Y. et al., 1997, Biol. Psychiatry, **42** : 898-903.

- Fraley et al., 1979, Proc. Natl. Acad. Sci. USA, **76** : 3348-3352.
- Fraley et al., J.Biol.Chem. 255 (1980) 10431
- Francis et al., 1995, J. Clin. Invest. , **1** : 78-87
- Fromont-Racine et al, 1993. Nucleic Acid Res.21(7) :1683-1684
- 5 Fuller S.A. et al., 1996, *Immunology in Current Protocols in Molecular Biology*, Ausubel et al.
- Gopal, 1985, Mol. Cell. Biol., **5** : 1188-1190.
- Graham et al., 1973, Virology, **52** : 456-457.
- Graham et al., J. Gen. Virol. **36** (1977) 59
- 10 Haff L.A. and Smirnov I.P., Genome Research, 7:378-388, 1997
- Ham, Methods Cell.Biol. 21a (1980) 255
- Hames BD and Higgins SJ, 1985, "Nucleic acid hybridization : a practical approach", Hames and Higgins Ed., IRL Press, Oxford.
- Hammeling et al., in: Monoclonal Antibodies and T-Cell Hybridomas,
- 15 Elsevier, N.Y., pp. 563-681 51981
- Harland et al., 1985, J. Cell. Biol., **101** : 1094-1095.
- Hoeg PNAS 93 p11448, 1996
- Horvath S. et al., 1998, Am. J. Hum. Genet., 1998, **63** : 1886-1897.
- Houben Weyl, 1974, in Methode der Organischen Chemie, E. Wunsch Ed.,
- 20 Volume 15-I et 15-II,
- Huygen et al., 1996, Nature Medicine, 2(8):893-898
- Kaneda et al., Science 243 (1989) 375
- Kim et al. Genomics (1996), **34** :213-218)
- Koch Y., 1977, Biochem. Biophys. Res. Commun., **74**:488-491
- 25 Köhler et al, Eur. J. Immunol. **6** :511 (1976)
- Köhler et al, Eur. J. Immunol. **6**:292 (1976)
- Köhler et al, Nature **256** :495 (1975)
- Kohler G. and Milstein C., 1975, Nature, **256** : 495.

- Kozbor et al., 1983, *Hybridoma*, 2(1):7-16.
- Lander and Schork, *Science*, 265, 2037-2048, 1994
- Langmann T. et al., 1999, *Biochem. Biophys. Res. Comm.*, 257 : 29-33.
- Leger OJ, et al., 1997, *Hum Antibodies*, 8(1): 3-16
- 5   Levrero et al., *Gene* 101 (1991)
- Luciani M.F. et al., 1994, *Genomics*, 21 : 150-159.
- Lyer V. et al., 1999, *Science*, 283 : 83-87.
- MackKinnon et al *Gene* 19 p33 1982
- Maniatis et al. *Molecular cloning : A Laboratory Manual*, Cold Spring Harbor
- 10   Press, Cold Spring, NY 1982
- Marcil M. et al., 1999, *Arteriosclerosis Throbosis and Vascular Biology*, 17 : 1813-1821.
- Martineau P, Jones P, Winter G, 1998, *J Mol Biol*, 280(1):117-127
- McCormick, *BioTechnology* 3 (1985) 689
- 15   McGrory et al (*Virology*, 163, 614,1988
- McLaughlin BA et al., 1996, *Am. J. Hum. Genet.*, 59 : 561-569.
- Merrifield RB, 1965a, *Nature*, 207(996): 522-523.
- Merrifield RB., 1965b, *Science*, 150(693): 178-185.
- Morrison et al., EP 173494
- 20   Morrison, *Science* 229 :1202 (1985)
- Narang SA, Hsiung HM, Brousseau R, *Methods Enzymol* 1979;68:90-98
- Neuberger et al., *Nature* 314 : 268 (1985)
- Neuberger et al., WO 8601533
- Nickerson D.A. et al., *Genomics*, 1992, 12 : 377-387.
- 25   Nicolau C. et al., 1987, *Methods Enzymol.*, 149:157-76.
- Oi et al., *Biotechnique* 4 :214 (1986)
- Okayama (*Mol Cell. Biol.* 7: 2745, 1987

- Pagano et al., J. Virol. 1 (1967) 891
- Potter et al., 1984, Proc Natl Acad Sci U S A., 81(22):7161-5
- Reimann KA, et al., 1997, AIDS Res Hum Retroviruses. 13(11): 933-943
- Remaley et al. ATVB 17,1813,1997 Chen et al Mol Cell Biol 7 p2745 1987
- 5 Ridder R, Schmitz R, Legay F, Gram H, 1995, Biotechnology (N Y), 13(3):255-260
- Rigaud et al. Biochimica et Biophysica Acta 1231 (1995) 223-246
- Robinson et al., WO 8702671
- Rust S. et al., Nature Genetics, vol. 20, Septembre 1998, pages 96-98
- 10 Sambrook, J. Fritsch, E. F., and T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2ed. Cold Spring Harbor Laboratory, Cold spring Harbor, New York.
- Samulski et al., 1989, J. Virol., 63 : 3822-3828.
- Sanchez-Pescador R., 1988, J. Clin. Microbiol., 26(10):1934-1938
- 15 Sham P.C. et al., Ann. Hum. Genet., 1995, 59 : 323-336.
- Sternberg N.L., 1992, Trends Genet., 8 : 1-16.
- Sternberg N.L., 1994, Mamm. Genome, 5 : 397-404.
- Strautnieks S.S. et al., 1998, Nature Genetics, 20 : 233-235.
- Tacson et al., 1996, Nature Medicine, 2(8):888-892.
- 20 Taniguchi et al., EP 171496
- Tur-Kaspa et al, 1986, Mol. Cell. Biol., 6 : 716-718.
- Urdea M.S., 1988, Nucleic Acids Research, 11: 4937-4957
- Urdea MS et al., 1991, Nucleic Acids Symp Ser., 24: 197-200.
- Vaisman JBC 270 p12269, 1995
- 25 Van, den Hazel. H., H. Pichler, V. M. M. do, E. Leitner, A. Goffeau, and G. Daum. 1999. PDR16 and PDR17, two homologous genes of *Saccharomyces cerevisiae*, affect lipid biosynthesis and resistance to multiple drugs. J Biol Chem. 274 (4):1934-41.

Wands et al., Gastroentérolology 80:225-232 (1981)

Xiong M. et al., 1999, Am. J. Hum. Genet., 64 : 629-640.

Yamon et al. Blood vol 90, n°8 pages 2911-2915, 1997

**REVENDICATIONS**

- 5 1. Acide nucléique comprenant au moins 245 nucléotides consécutifs d'un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 1-14, ou un acide nucléique de séquence complémentaire.
2. Acide nucléique comprenant un polynucléotide choisi dans le groupe  
10 constitué des séquences nucléotidiques SEQ ID NO 15-47 , ou un acide nucléique de séquence complémentaire.
3. Acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques  
15 SEQ ID NO 48-90, ou un acide nucléique de séquence complémentaire.
4. Acide nucléique ayant au moins 80% d'identité en nucléotides avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 48-90, ou un acide nucléique de séquence complémentaire.  
20
5. Acide nucléique hybridant, dans des conditions de forte stringence, avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 48-90, ou un acide nucléique de séquence complémentaire.  
25
6. Acide nucléique comprenant un polynucléotide de séquence SEQ ID NO 91, ou un acide nucléique de séquence complémentaire.
7. Acide nucléique comprenant au moins huit nucléotides consécutifs d'un  
30 polynucléotide choisi dans le groupe constitué des séquences nucléotidiques



SEQ ID NO 92, un fragment biologiquement actif de celui-ci ou un acide nucléique de séquence complémentaire.

8. Acide nucléique ayant au moins 80% d'identité en nucléotides avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 92, un fragment biologiquement actif de celui-ci ou un acide nucléique de séquence complémentaire.

9. Acide nucléique hybridant, dans des conditions de forte stringence, avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 92, un fragment biologiquement actif de celui-ci ou un acide nucléique de séquence complémentaire.

10. Acide nucléique ayant au moins huit nucléotides consécutifs d'un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 93-96, ou un acide nucléique de séquence complémentaire.

11. Acide nucléique ayant au moins 80% d'identité en nucléotides avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 93-96, ou un acide nucléique de séquence complémentaire.

12. Acide nucléique hybridant, dans des conditions de forte stringence, avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 93-96, ou un acide nucléique de séquence complémentaire.

13. Acide nucléique ayant au moins huit nucléotides consécutifs d'un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 97-108 et comprenant la base polymorphe, ou un acide nucléique de séquence complémentaire.

14. Sonde ou amorce nucléotidique spécifique du gène ABC1, d'une longueur d'au moins 15 nucléotides choisie parmi les acides nucléiques selon l'une quelconque des revendications 1 à 9.
- 5
15. Sonde ou amorce selon la revendication 14, comprenant un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 109-138, ou un acide nucléique de séquence complémentaire.
- 10
16. Sonde ou amorce nucléotidique utile pour la détection d'une mutation dans le gène ABC1, d'une longueur d'au moins 15 nucléotides, choisie parmi les acides nucléiques selon l'une des revendications 11 et 12.
- 15
17. Sonde ou amorce nucléotidique selon la revendication 16, comprenant un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID NO 109 – 112, ou un acide nucléique de séquence complémentaire.
- 20
18. Sonde ou amorce nucléotidique utile pour la détection d'un polymorphisme dans le gène ABC1, d'une longueur d'au moins 15 nucléotides, choisie parmi les acides nucléiques selon la revendication 13.
- 25
19. Sonde ou amorce selon la revendication 18, comprenant un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID NO 142-149, ou un acide nucléique de séquence complémentaire.
- 30
20. Amorce nucléotidique comprenant au moins 15 nucléotides consécutifs d'un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 97-108 ou de leurs séquences complémentaires, la base de l'extrémité 3' de ces amorces étant complémentaire du nucléotide localisée immédiatement du côté 5' de la base polymorphe de l'une des séquences SEQ ID NO 97-108 ou de leurs séquences complémentaires.

21. Amorce nucléotidique comprenant au moins 15 nucléotides consécutifs d'un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 97-108 ou de leurs séquences complémentaires, la base de l'extrémité 3' de ces amorces étant complémentaire d'un nucléotide situé à 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 nucléotides ou plus du côté 5' de la base polymorphe de l'une des séquences SEQ ID NO 97-108 ou de leurs séquences complémentaires.
22. Procédé pour l'amplification d'un acide nucléique selon l'une quelconque des revendications 1 à 13 contenu dans un échantillon, ladite méthode comprenant les étapes de :
- a) mise en contact de l'échantillon dans lequel la présence de l'acide nucléique cible est suspectée avec une paire d'amorces nucléotidiques dont la position d'hybridation est localisée respectivement du côté 5' et du côté 3' de la région de l'acide nucléique cible dont l'amplification est recherchée, en présence des réactifs nécessaires à la réaction d'amplification; et
  - b) détection des acides nucléiques amplifiés.
23. Procédé d'amplification selon la revendication 22, caractérisé en ce que les amorces nucléotidiques sont choisies parmi les amorces selon l'une quelconque des revendications 14 à 19.
24. Nécessaire pour l'amplification d'un acide nucléique selon l'une quelconque des revendications 1 à 13 comprenant :
- a) un couple d'amorces nucléotidiques dont la position d'hybridation est localisée respectivement du côté 5' et du côté 3' de l'acide nucléique cible dont l'amplification est recherchée;
  - b) le cas échéant, les réactifs nécessaires à la réaction d'amplification.

25. Nécessaire pour l'amplification d'un acide nucléique selon la revendication 22, caractérisé en ce que les amorces nucléotidiques sont choisies dans le groupe constitué des amorces selon l'une des  
5 revendications 14 à 19.

26. Sonde nucléotidique selon l'une quelconque des revendications 14 à 19, caractérisée en ce qu'elle comprend un composé marqueur dont la présence est détectable.

10

27. Procédé de détection de la présence d'un acide nucléique selon l'une quelconque des revendications 1 à 13 dans un échantillon, ladite méthode comprenant les étapes de :

a) mettre en contact une ou plusieurs sondes nucléiques selon l'une  
15 des revendications 14 à 19 avec l'échantillon à tester;

b) détecter le complexe éventuellement formé entre la ou les sondes et l'acide nucléique présent dans l'échantillon.

28. Procédé de détection selon la revendication 27, caractérisé en ce que la  
20 ou les sondes sont immobilisées sur un support.

29. Nécessaire pour la détection de la présence d'un acide nucléique selon l'une quelconque des revendications 1 à 13 dans un échantillon, ledit nécessaire comprenant :

a) une ou plusieurs sondes nucléotidiques selon l'une quelconque des  
25 revendications 14 à 19;

b) le cas échéant, les réactifs nécessaires à la réaction d'hybridation.

30. Nécessaire de détection selon la revendication 29, caractérisé en ce que  
30 la ou les sondes sont immobilisées sur un support.

31. Vecteur recombinant comprenant un acide nucléique selon l'une des revendications 1 à 13.

32 Vecteur selon la revendication 31 caractérisé en ce qu'il est un  
5 adénovirus.

33 Vecteur selon l'une des revendications 32 et 33 caractérisé en ce qu'il est l'ABC1-rIdV

10 34 Cellule hôte recombinante comprenant un acide nucléique selon l'une des revendications 1 à 13 ou un vecteur recombinant selon l'une des revendications 31 à 33.

35. Polypeptide ABC1 muté, caractérisé en ce qu'il comprend un polypeptide  
15 de séquence d'acides aminés SEQ ID NO 140.

36. Polypeptide ABC1 muté, caractérisé en ce qu'il comprend un polypeptide de séquence d'acides aminés SEQ ID NO 141.

20 37. Anticorps dirigé contre un polypeptide ABC1 muté selon l'une quelconque des revendications 35 et 36, ou un fragment peptidique de ce dernier..

38. Anticorps selon la revendication 37, caractérisé en ce qu'il comprend un  
25 composé détectable.

39. Procédé pour détecter la présence d'un polypeptide selon l'une des revendications 35 ou 36 dans un échantillon, comprenant les étapes de:

a) mise en contact de l'échantillon avec un anticorps selon l'une des  
30 revendications 37 ou 38;

b) détection du complexe antigène/anticorps formé .

40. Nécessaire de diagnostic pour la détection de la présence d'un polypeptide selon l'une des revendications 35 ou 36 dans un échantillon,  
35 ledit nécessaire comprenant:

- a) un anticorps selon l'une des revendications 37 ou 38;
- b) un réactif permettant la détection des complexes antigènes/anticorps formés.

- 5 41. Composition pharmaceutique destinée à la prévention ou au traitement de sujets affectés d'un dysfonctionnement du transport inverse du cholestérol, comprenant un acide nucléique selon l'une des revendications 1 et 6, en association avec un ou plusieurs excipients physiologiquement compatibles.
- 10 42. Composition pharmaceutique destinée à la prévention ou au traitement de sujets affectés d'un dysfonctionnement du transport inverse du cholestérol, comprenant un vecteur recombinant selon la revendication 31, en association avec un ou plusieurs excipients physiologiquement
- 15 compatibles.
43. Utilisation d'un acide nucléique selon l'une des revendications 1 et 6 pour la fabrication d'un médicament destiné à la prévention de l'Athérosclérose sous diverses formes ou plus particulièrement au traitement
- 20 de sujets affectés d'un dysfonctionnement du transport inverse du cholestérol.
44. Utilisation d'un vecteur recombinant selon la revendication 31 pour la fabrication d'un médicament destiné à la prévention de l'Athérosclérose sous
- 25 diverses formes ou plus particulièrement au traitement de sujets affectés d'un dysfonctionnement du transport inverse du cholestérol.
- 45 Utilisation selon la revendication 44 caractérisé en ce que le vecteur est l'ABC1-rdV.
- 30 46. Utilisation du polypeptide ABC1 de séquence SEQ ID NO 139 pour la fabrication d'un médicament destiné à la prévention de l'Athérosclérose sous diverses formes ou plus particulièrement au traitement de sujets affectés d'un dysfonctionnement du transport inverse du cholestérol.

47. Composition pharmaceutique pour la prévention ou le traitement de sujets affectés d'un dysfonctionnement du transport inverse du cholestérol, comprenant une quantité thérapeutiquement efficace du polypeptide de séquence SEQ ID NO 139.

5

48 Utilisation du polypeptide ABC1, ou de cellules exprimant le polypeptide ABC1, pour cribler des principes actifs pour la prévention ou le traitement de maladies résultant d'un dysfonctionnement du transport inverse du cholestérol.,

10

49. Procédé de criblage d'un composé actif sur le métabolisme du cholestérol, agoniste ou antagoniste du polypeptide ABC1, ledit procédé comprenant les étapes suivantes :

- a) préparer des vésicules membranaires contenant le polypeptide ABC1 et un substrat lipidique comprenant un marqueur détectable ;
- b) incuber les vésicules obtenues à l'étape a) avec un composé candidat agoniste ou antagoniste ;
- c) mesurer qualitativement et/ou quantitativement la libération du substrat lipidique comprenant un marqueur détectable ;
- d) comparer la mesure obtenue à l'étape b) avec une mesure de la libération du substrat lipidique marqué par les vésicules n'ayant pas préalablement été incubées avec le composé candidat agoniste ou antagoniste.

50. Procédé de criblage d'un composé actif sur le métabolisme du cholestérol, agoniste ou antagoniste du polypeptide ABC1, ledit procédé comprenant les étapes suivantes :

- a) obtenir des cellules, par exemple une lignée cellulaire, exprimant naturellement ou après transfection le polypeptide ABC1 ;
- b) incuber les cellules de l'étape a) en présence d'anion marqué par un marqueur détectable ;
- c) laver les cellules de l'étape b) afin d'éliminer l'excès de l'anion marqué n'ayant pas pénétré dans ces cellules ;
- d) incuber les cellules obtenues à l'étape c) avec un composé candidat agoniste ou antagoniste du polypeptide ABC1 ;
- e) mesurer l'efflux de l'anion marqué ;

f) comparer la valeur de l'efflux de l'anion marqué déterminé à l'étape e) avec la valeur de l'efflux de l'anion marqué mesuré avec des cellules n'ayant pas préalablement été incubées en présence du composé candidat agoniste ou antagoniste du polypeptide ABC1.

5

51. Procédé de criblage d'un composé actif sur le métabolisme du cholestérol, agoniste ou antagoniste du polypeptide ABC1, ledit procédé comprenant les étapes suivantes :

- a) cultiver des cellules d'une lignée monocyttaire humaine dans un milieu de culture approprié, en présence d'albumine humaine purifiée ;
- 10 b) incuber les cellules de l'étape a) simultanément en présence d'un composé stimulant la production d'IL-1 beta et du composé candidat agoniste ou antagoniste;
- c) incuber les cellules obtenues à l'étape b) en présence d'une concentration
- 15 appropriée d'ATP ;
- d) mesurer d'IL-1 beta libérée dans le surnageant de culture cellulaire.
- e) comparer la valeur de la libération de l'IL-1 beta obtenue à l'étape d) à la valeur de l'IL-1 beta libérée dans le surnageant de culture de cellules n'ayant pas été préalablement incubées en présence du composé candidat agonsite
- 20 ou antagoniste.



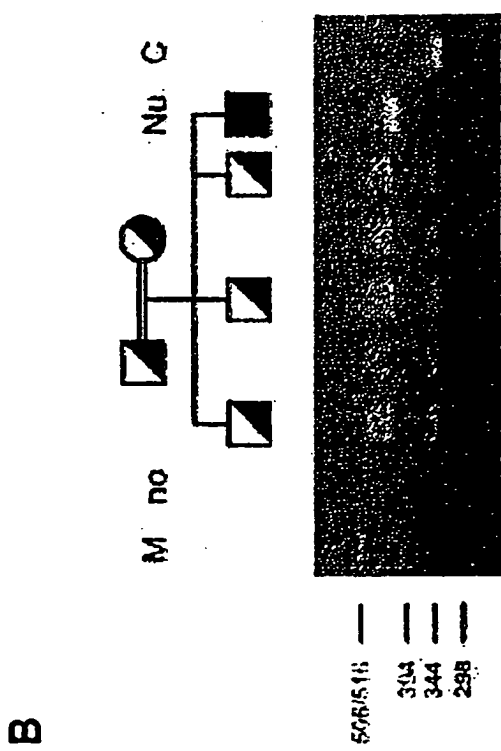
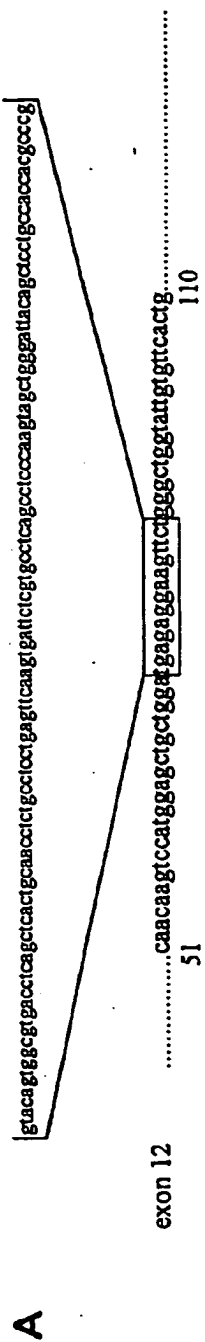


Figure 1

2/2

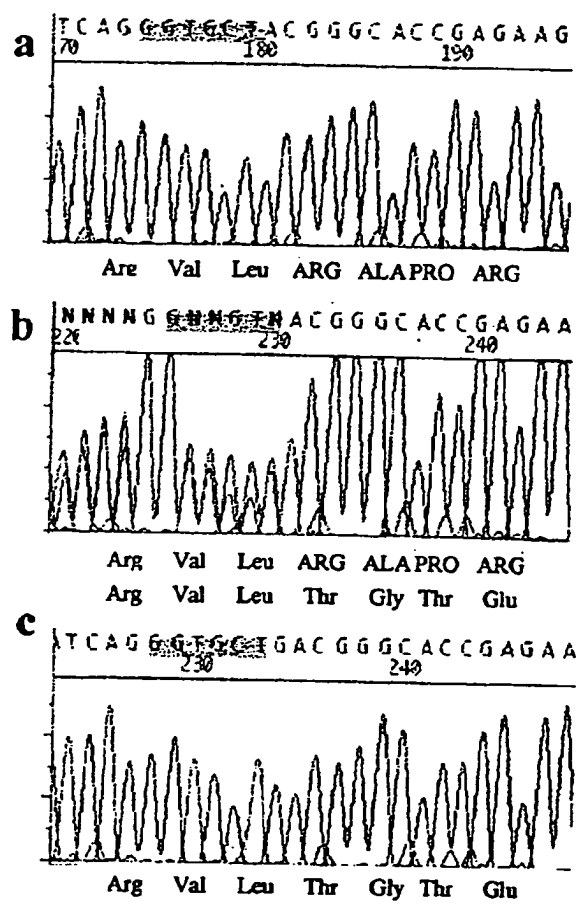


Figure 2

## LISTE DE SEQUENCES

- <110> Rhône Poulenc Rorer S.A.
- 5 <120> Acides nucléiques et protéines du gène ABC1 humain et leur application en thérapie et diagnostic
- <130> RPR-ABC1-1
- 10 <140>  
<141>
- <160> 141
- 15 <170> PatentIn Ver. 2.1
- <210> 1  
<211> 3153  
<212> ADN
- 20 <213> Homo sapiens
- <400> 1  
 acatggcaat ggcattcatt aggaatctag ctgggaaaaat ccagtgtgta tgcttggaaa 60  
 tgagggatct ggggctggag agaaaggcat gggcatgcct tggagggact tgtgtgtcaa  
 120  
 gctgaggacc ttacttttaa gctctagggg accaggcaag gggagatgta gatacgttac  
 180  
 tctgatgggg tggatgaatt gaagaaggat gaggcaagaa tgaaggcaga gaccagggag  
 240  
 gaggtctctc aagtggccaa ggcataaagc aagaaatgag gcctgggtgac tgcttagtgg  
 300  
 cagagcagtg aaagagaggg aggcatacaa gtgagtctcg atttctagct ggggtgggtg  
 360  
 tagcgatgtc cagtaggcca gtggctactg aggtctgcag tggaggaggg tggttgggct  
 420  
 ggagacagat gatgagggag tcatcagcct gtgggtggaa gaaaaggga cctcttccaa  
 480  
 ctgttttctt tgcttcttcc ctctctttct cttttttttt ttttttggac agagtcttgc  
 540  
 tctgtcacc aggctgaaat gcagtggcat gatcttggct caccacagcc tccgcctcct  
 600  
 gggttcaagc aattctctcg tctcagcctc cagagtagct gggattacag gcacatatca  
 660  
 ctgtgcccg ctaatttttg tattttcagt ggagatggga ttccaccatg ttggtcgggc  
 720  
 tggaatgaac tcttgacctc aagtgatcca cctgcctcag cctcccaaag tgttgggatt  
 780  
 acaggcattg agccaccgag ccggccttt cttccctctc ttaaagagtg tttatttaat  
 840  
 tccacaaaca tgagcttgtc acccctgta gcctggcacc tccacacga ggtgatggct  
 900  
 gaggcttctg cttctgctgg ggtagctctg atctttctgc tttctctggc actgtctacc  
 960  
 catgttgctt caccacacag gtcccagggc acctctctcg ggcaagtctt ggaacctctt  
 1020  
 gacactgatt tgctctcttt tctgagctgc ttttagccac ccatcctcgg gacctgtttt  
 1080  
 ctctctgctt ccacctctgc gggcagctct aggtctcctg cccctcacga gcacctcaga  
 1140  
 gaggccacgt gctcagtgat ctccagtggc gcactcttct agtcttgcta ttctttttgg  
 1200  
 ccatgttggt cagaaccat actgggcagg gccgacttca ccctaaaggc tgcgtctctt  
 1260  
 cactctgctt ttgtttgttc caataaagt ggcttcagaa ttgctaacce tagcctctgt  
 1320  
 gaacttgatg ggtacaattt tgtgtctgtt atgttaacaa aaatacatat ataccttctt  
 1380

ggtgatggta taaattgcta ttctctattg gaaagcaatt tggaaatgaaa atttaaagaa  
1440  
ccattttaaa atatgctatc ctgcgctacct ccattccacc caccceccagg gatgtagcct  
1500  
5 actgaaataa ttttaaagaa gtcaccatat gagagaaaat gttattgcta tattgttatt  
1560  
gtgagaaatt ggaaatagac taaatgttca gcactatagg aataattaat gaaattacat  
1620  
atctctata caatcattat gctgccattg aaataataaa taaaaaggcg caagggggga  
1680  
10 aaagcttata atgttagtga aactaagact gattttttta taaagcagca gttttcagac  
1740  
ccttggagac tccaattcgg tagaaccaga gcttcactct ctctgtcgaa gctgtgacag  
1800  
15 gagttgcaa tgccctctct ttttgcctgag ttgcagctg ctgtttttcc ggcagcacat  
1860  
ctgtgcaggc ctctgcctcg gccctctggt atctgctgat tgagcagcgg attgatctgt  
1920  
ccttctcttt cgtgttgacc catgtgagga accaactggc aaggggaacaa gaaatggaaa  
1980  
20 taggcctctt ttgcatcatg acctgtacat cctgcaattg gaaaagattg tactttagtt  
2040  
ggtttaacca gcagcattat ttttctaacc taagcagtaa gaaggaatta ggttttatgt  
2100  
25 ggatcaaca gactgggtct caaaagagga aggtgataga acacagtggg gagggggagg  
2160  
tgactagaa acagagggcc tatgctttca ttctggcttt gctacttaat agctgtgtga  
2220  
cccaatctta gagacttaac ctctctgaac ttccattttc tcatgtataa aatgggaaat  
2280  
30 attaaaggat actcactggg ctggtggcct gtgcctgtaa tcccagcact tggggagggt  
2340  
gaggtgggag gatcacttga gcccagggtg tcaagaccag cccaggcaac atggcaagac  
2400  
35 tctgtctcta tgaaaaaatt aaaaattagc caggtgtggt ggtgtgcacc tgtagtctta  
2460  
gctacttggg aggtgagat gggaggatca ctggggcttg ggaggtcaag gctgcggtga  
2520  
gctgtgatcc catcactgca ctccagcccc ggccggcagag cgagacactg aatccaaacg  
2580  
40 acaacaacaa caaaaggcaa aaaaataaaa gtgccctctt tatggagttg tgtaagggtga  
2640  
agcatatata ctattcaaca tagtaactat ataaaggaag tattgttgtt gttactgtag  
2700  
45 ttaataccat taagtgagat gtttcgtata gtggaaagca catggactct gaattcagac  
2760  
tggcttgact ttgagtctca gctccacatc tagtaatact atgaccaagc cctgggttaa  
2820  
50 atcatgtttt tttttcttca gccttagtct tctcacatat aaaatagggg cactgtcatt  
2880  
tacctcagtt ttctgtgagg ataaaacaac gacagtgtat atgcaagtat tttgtaaatt  
2940  
ttgtagtgtt cctcaagatt tagttgggtt ttactacttg tactttctca ctggaatggc  
3000  
55 agatgctgtt ggacagcagg gacaatgacc acttttggga acagcagttg gatggcttag  
3060  
attggacagc ccaagacatc gtggcggttt tggccaagca cccagaggat gtccagtcca  
3120  
gtaatgggtt tgtgtacacc tggagagaag ctt  
60 3153

<210> 2  
<211> 7660  
65 <212> ADN  
<213> Homo sapiens  
  
<400> 2

aattcgggtc caattaaatt ttgaaattt tatattaaa attatattag tagggatggg 60  
taagagggtg ttgtgtctgg ttggttggtt agttgctatg actcagaatt gctaagaaaa  
120  
cagaaaagta agataagatc attgttttaa cctcttttcc tccacaaaat caataaataa  
5 180  
catatcccta aattactctt agaatttctc ttaaattgca gtgaaaaacc aaaatccttc  
240  
attcttggtt gaaggttgga aaactacgtt agagaggatt agagagagag gatgagcaat  
300  
10 cgtgtagtca gcccttgccct cctagtgtag gatttgtctc agccactgct tgttgtctcg  
360  
gctgccaaag ttctcatgaa ggctgttctt ctatcagtgt gtcaacctga acaagctaga  
420  
acccatagca acagaagtct ggctcatcaa caagtccatg gagctgctgg atgagaggaa  
15 480  
gttctgggct ggtatttgtt tcaactggaat tactccaggc agcattgagc tgcccatca  
540  
tgtcaagtac aagatccgaa tggacattga caatgtggag aggacaaata aaatcaagga  
600  
20 tgggtaagtg gaatcccatc acaccagcct ggtcttgggg aggtccagag cacctattat  
660  
attaggacaa gaggtacttt attttaacta aaaatttggg agaaatttca acaacaacaa  
720  
aaaaactcaa ctgtgtgtca tgattttggt gaaattggta catgacttgc tggagggttt  
25 780  
ttcataggtc ataaaataac agtatctttt gatttagcat ttctactcaa gggaattaat  
840  
tccaggaatt ttggtggcag gcacctgtaa tcccagctac tcgggaggct gaggcaggag  
900  
30 aattgcttga acccaggagg cagagggtgc agtgagctaa gatcgcatca ttgcaactcc  
960  
gcctgggcaa taagagtga actccatctc aaaaaaaaaa aaagatacaa aaatagaaaa  
1020  
aggggcttgg taagggtagt agggttttgg gcaatttttt tttttttttt tttttattgt  
35 1080  
atggttctaa aggaatggtt gattacctgt ggtttggttt taggtactgg gacctggtc  
1140  
ctcgagctga cccctttgag gacatgcggt acgtctgggg gggcttcgcc tacttgagg  
1200  
40 atgtggtgga gcaggcaatc atcagggtgc tgacgggcac cgagaagaaa actggtgtct  
1260  
atatgcaaca gatgccctat ccctgttacg ttgatgacat gtaagttacc tgcaagccac  
1320  
tgtttttaac cagtttatac tgtgccagat gggggtgtat atatgtgtgt gcatgtgcat  
45 1380  
gcatgtgtga atgatctgga aataagatgc cagatgtaag ttgtcaacag ttgcagccac  
1440  
atgacagaca tagatatatg tgcacacact agtaaacctc tttccttctc atccatgggt  
1500  
50 gccactttta tctttttatt tttatttttt tttttgagat ggagtctcgc tctgacgccc  
1560  
aggctggagt gcagtggctc gatctcggtt cactgcaacc tttgcctccc gggttcaagg  
1620  
tattctctg cctcagcctc cacagtagct gggactacag gctcatgctg ccacgcccgg  
55 1680  
ctgaactttt gtattttagt agagacgagg ttccaccatg ttaccaggc tagacttcaa  
1740  
ctcctgagct caggcaatcc accctccttg gcctcccaaa gtgctgggat tacaggtgtg  
1800  
60 agccactgca ccagagccac cactttaatt ttttactc tacccttttg gtcaaaattt  
1860  
gtcaatctg caagcttaaa atgtgtcatg acaaacacat gcaagcacat actcacacat  
1920  
agatgcagaa acagcgtcta aacttataaa agcacagttt atgtaaatgt gtgcacttct  
65 1980  
tctccctagg tggtaaacca cttttcaaaa caaccctaat aaaactgaac aaagcttctt  
2040

cctcttagac tttttagaaa atctttcagt gctgagtcac taagctgcca agttctcatt  
2100  
gtgggaacta tgcctttgga tgtaatgatt tcttctaaga caatgggcgg aggtgtagtt  
2160  
5 attgcagaca tctgaaatat gtaatgttcc tccagattc tggaaattct cttattctct  
2220  
gtggttggtg gtggtggtg gatgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg  
2280  
tagggatcag gatgcgggag gagctgggtt ctgcttgat tggttctctg ttttgcattg  
10 2340  
aatagtgtgt ttccttgat ggctatctat agcttttcaa ggtcaccaga aattatcctg  
2400  
ttttcacct tctaaacaat tagctggaat ttttcaaagg aagactttta caaagacccc  
2460  
15 taagctaagg tttactctag aaaggatgtc ttaagacagg gcacaggagt tcagaggcat  
2520  
taagagctgg tgcctgttgt catgtagtga gtatgtgcct acatggtaaa gctttgacgt  
2580  
gaacctcaag ttcagggtcc aaaatctgtg tgccttttta ctttgacat ctgcattttc  
20 2640  
tattctagct tggaaatctga aacattgaca agagctgcct gaaatgtatg tctgtggtgt  
2700  
gattagagtt acgataagca agtcaatagt gagatgacct tggagatgtt gaacttttgt  
2760  
25 gagagaatga gttgtttttt tgttttggtt tttagtactt taacataatc tacctttagt  
2820  
ttaagtatcg ctacacagtt cctagttact gaagcaagcc cccaaagaaa tttggtttgg  
2880  
caacactttg ttagcctcgt ttttctctct acattgcatt gctcgtgaag cattggatca  
30 2940  
tacgtacatt tcagagtcta gagggcctgt ccttctgtgg ccagatgtg gtgctccctc  
3000  
tagcatgcag gctcagagcc cttggcccat caccctggct cactgtgtgc tttctttctc  
3060  
35 cccttgctct tccttggggc ctccagcttt ctgcggtgta tgagccggtc aatgcccctc  
3120  
ttcatgacgc tggcctggat ttactcagtg gctgtgatca tcaagggcat cgtgtatgag  
3180  
aaggaggcac ggctgaaaga gaccatgcgg atcatgggcc tggacaacag catcctctgg  
40 3240  
tttagctggt tcattagtag cctcattcct cttcttgtga gcgctggcct gctagtgtgc  
3300  
atcctgaagg taaggcagcc tctctgctc ttcctgcca ggaaactccg aaatagctca  
3360  
45 acacgggcta agggaggaga agaagaaaaa aaatccaagc ctctggtaga gaaggggtca  
3420  
tacctgtcat ttctgcaat ttcatccatt tatagttggg gaaagtgagg ccagagagg  
3480  
ggcagtgact tgcccaaggc caaccagcc gggtagcagc taagtaggat gagagtgcag  
50 3540  
ggttcatgct ttccagataa ccacatgctc aactgtgcca tgcgtctca ttggtagtgg  
3600  
ttcatggcag catctgaaag ctatttattt tcttagatat attgggtggc gattcttctc  
3660  
55 aagtttctaa gaacaataat cagaaggata tatattgttg caggtagac tgtctggaag  
3720  
cagaggctga aatagagttt gatgtatggg tatttatgag ggctcaatac ctatgaagag  
3780  
atatggaaga tgcaggattg ggcagaggga ggagtgaac tgtgatatag ggccaacccc  
60 3840  
gtggggcact ctanagaata tgcagcttgt tggagttgtt ntcatcgag ctgaaacatc  
3900  
cagccctttg tgctcccca aggcctccct cctgacacca cctacctcag ccctctcaat  
3960  
65 caatcactgg atgtgggctg ccctgggaag gtcgtgcccc agggcctaca tggctctctg  
4020  
ctgctgtgac aaaccagag ttgctgatgc ctgaggccgt ctactgacag ctgggcaaca  
4080

aggcctccct gaatggggac tctgggcagt gcagttttgt gtctgaacca tacattaata  
 4140  
 tatttatatc cgaattttct ttctctgcaa gcatttcata taaagacaca tcaggtaaaa  
 4200  
 5 ataatgttt ttgaagcaaa aggagtacaa agagataaga actaactaat ttaatactag  
 4260  
 ttaccatctg ttacaaatag ttctactga ttgccaagga ctgtttaaac acatcacatg  
 4320  
 10 ggcttcttct tctatcctca ctaacccttt taacagacaa ggaaatgagg ctcaggaagg  
 4380  
 tcaaggactt tattgaggtt ccacagtagg atacagttct tgctaaaagc aaccctcccc  
 4440  
 tcatgctctg ttatctaact gcaaggggaa ggtcagtggc agaggtagtg gtcccatggt  
 4500  
 15 tggcgcataa gagctgctct gagacaactg catgctggtg ggtcctgcag acatgtacce  
 4560  
 atcagccgga gataggctca aaatatccac aagagtttgg atgattgtgg gaatgcagaa  
 4620  
 tccatggtga tcaagagggg aagtcaagtt gcctggccat ttcccttggc ttttagacag  
 4680  
 20 aaaagttacg tgggatatta tctccacag ctcttctgtg gtgccaccag tcatagtctt  
 4740  
 tatataagga gaaaccagtt gaaattacct attgaagaaa caaagagcaa actcgccac  
 4800  
 25 tgaatgcgt agaaagccct ggactctgtt gtattcataa ctctgccatt atttttctgc  
 4860  
 gtagttttgg gtaagtcact tatcttcttt aggatggtta tgatcagttg cctcatcaga  
 4920  
 aagatgaaca gcattacgcc tctgcattgt ctctaactg agtaggaata aaccctgtct  
 4980  
 30 tttttctgta gatcatacaa gtgagtgtt gggattgtt aggcagcaca ttgatgtgt  
 5040  
 ctcttcttc ccagttagga aacctgctgc cctacagtga tcccagcgtg gtgtttgtct  
 5100  
 35 tcctgtccgt gtttctgtg gtgacaatcc tgcagtgtt cctgattagc acactcttct  
 5160  
 ccagagccaa cctggcagca gcctgtggg gcacatcta ctacagctg tacctgccct  
 5220  
 acgtcctgtg tgtggcatgg caggactacg tgggcttca actcaagatc ttcgctgtga  
 5280  
 40 gtacctctgg ctttcttca gtggctgtag gcatttgacc ttcctttgga gtcctgaat  
 5340  
 aaaagcagca agttgagaac agaagatgat tgtcttttcc aatgggacat gaaccttagc  
 5400  
 45 tctagattct aagctcttta aggtaaggg caagcattgt gttttattaa attgtttacc  
 5460  
 tttagtcttc tcagtgaatc ctggttgaat tgaattgaat ggaatttttc cgagagccag  
 5520  
 actgcattct gaactgggct ggggataaat ggcattgagg aatggcttca ggcaacagat  
 5580  
 50 gccatctctg ccctttatct cccagctctg ttggctatgt taagctcatg acaaagccaa  
 5640  
 ggccacaaat agaactgaaa actcttgatg tcagagatga cctctcttgt cttccttgtg  
 5700  
 55 tccagtatgg tgttttgcct gagtaatgtt ttctgaacta agcacaactg aggagcaggt  
 5760  
 gcctcatccc acaaattcct gacttggaca ctctcttccc tcgtacagag cagggggata  
 5820  
 60 tcttggagag tgtgtgagcc cctacaagtg caagttgtca gatgtcccca ggtcacttat  
 5880  
 caggaaagct aagagtgact cataggatgc tcctgttgcc tcagtctggg cttcataggg  
 5940  
 atcagcagcc ccaaacaggc acctctgatc ctgagccatc cttggctgag cagggagcct  
 6000  
 65 cagaagactg tgggtatgcg catgtgtgtg ggggaacagg attgctgagc cttggggcat  
 6060  
 ctttggaaac ataaagtttt aaaagtttta tgcttactg tatatgcatt tctgaaatgt  
 6120

ttgtatataa tgagtgggta caaatggaat cattttatat gttacttggg agcccaccac  
 6180  
 tcccctaaag ggactctata ggtaaatact acttctgcac cttatgattg atccattttg  
 6240  
 5 caaattcaaa tttctccagg tataatttac actagaagag atagaaaaat gagactgacc  
 6300  
 aggaaatgga taggtgactt tgccctgttc tcacagagcc tgctgtctcc tgtggctttt  
 6360  
 10 gggtttggct gtgagtactt tgcccttttt gaggagcagg gcattggagt gcagtgggac  
 6420  
 aacctgtttg agagtccgtt ggaggaagat ggcttcaatc tcaccacttc ggtctccatg  
 6480  
 atgctgtttg acaccttcct ctatgggggtg atgacctggt acattgaggc tgtctttcca  
 6540  
 15 ggtacactgc tttgggcate tgtttgaaa atagtacttc tagctgatgt cctttctttg  
 6600  
 tgctagaatc tctgcagtgc atgggcttcc ctgggaagtg gtttgggcta tagatctata  
 6660  
 20 gtaaacagat agtccaagga caggcagctg atgctgaaag tacaattgtc actacttgta  
 6720  
 cagcacttgt ttcttgaaaa ctgtgtgcca ggcagcatgc aaaatgtttt atacacattg  
 6780  
 cttcatttaa ttctcacaag gctactctga agtagttact ataataacca gcaattttca  
 6840  
 25aatgagagaa ctgtgactca aagacgttaa gtaaccagct ttggtcacac aactgttaaa  
 6900  
 tgttggtagc tggaggtgaa tccacttcgg ttactctggg tcaataagcc caggcgaatc  
 6960  
 30ctcccaatgc tcaccaatt ctgtatttct gtgtcctcag aggggggtaca actaggagag  
 7020  
 gttctgtttc ctgagtacag gttgttaata attaaatata ctactctaa ggcctgcctg  
 7080  
 tgatttaatt agcattcaat aaaaattcat gttgaatttt tcttttagtac ttctttctta  
 7140  
 35atataatata tcttcttgac caagtccaag aggaacctgc gttggacagt ttcatatga  
 7200  
 gatcaaatc tgagagagca agatttaacc ctttttgggt caccctctga tccctcccta  
 7260  
 40aggaggtata catgaaatat ttattactcc tgccctgaact tctttcattg aatatgcaat  
 7320  
 tttgcagcat gcagattctg gatttaaat ctgagtctta acttactggc tgagggacct  
 7380  
 tggataggct ccttatccct cagtctctc atctctaaaa tggggatggc acctgccccg  
 7440  
 45tgggttgttg gaaggactta cagaggtgca gaatgtactg tgtacatagc aggtttcagc  
 7500  
 aaatgttagc tccctctttc cccacatcca ttcaaatctg ttccttctcc aaaggatgtg  
 7560  
 tcaaggagga aatggacctg gctgggaaac cctcagaata ctgggatgat gctgagcttg  
 7620  
 50gctcatacct gtgctttgct ttcaggccag tacggaattc  
 7660

55 <210> 3  
 <211> 1285  
 <212> ADN  
 <213> Homo sapiens

60 <400> 3  
 gaattcccag gccctgggat tttcttgc ccaagtccta ctggtttggc gaggaagtg 60  
 atgagaagag ccacctgggt tccaaccaga agagaatatc agaaagtaag tgctgttgac  
 120  
 ctctgtctct ttctttaacc tagtctgtct gctctgcta actgttgggg gcaagcgatg  
 180  
 65 tctcctgcct ttctaaaaga ctgtgaaacc actccagggg cagagaaatc acatgcagt  
 240



tccctttcca aatcctccca tgccatttat gtccaatgct gttgacctat tgggagttca  
 300  
 cgggtctcgat ccttgagga cttttctttt gttgtcttgg cttctagaag agtatctttt  
 360  
 5 acttgccccc tcccaaacac acatttcatg gtctcctaac aagctagaag aaagaggtaa  
 420  
 agacaagcgt gattgtgga ccatagcctc gctgcctgcc tgtgacatgg tgacctgtgt  
 480  
 atcagcctgt gtgggctgag accaagtggc taccacagag ctacagcctat gcttcataat  
 540  
 10 gtaatcatta cccagatccc taatcctctc ttggctctta actgcagaca gagatgtcca  
 600  
 cagctcatca aaggctctgc cttctgggtt ctttgtgctt agagtggctt cctaaatatt  
 660  
 15 taataggtec cttttctgcc agtctcttct gtgccatcc cctgattgcc cttggtaaaa  
 720  
 gtatgatgcc ccttagtgta gcacgcttgc ctgctgttcc taatcatctt ctctacctc  
 780  
 ctctttacac ctgactctg tttcagtcac ctagaaatgc tcacagtcgc tggaatatgt  
 840  
 20 catgttcttc cacacctcca tgcctttgta ggtactgttt gctctcacag gagaacttct  
 900  
 tctctaactt gcctatcttc tcaactcctc ctttctctcc aagatctagt tccggatccc  
 960  
 25 ctcccctgag catccctcct tggttctcag gtagtcagtc actctctgcc ctgaacttcc  
 1020  
 atggcacgtg aaagaaaatc tttttatttt aaaacaatta cagactcaca agaagtaata  
 1080  
 caaattacat gaggggggtc ccttaaacct ttcattccagt tccccaatg gtagcagcat  
 1140  
 30 gtgtaactgt agaatagtat caaaaccatg aaattgacat aggtacaatt cacaacctt  
 1200  
 cttcagattt cactagcttt atgtgcgctc atttgtgtgt gtgtgtgcgt atttagttct  
 1260  
 35 atgcaatttt atcatgtgtg aattc  
 1285

<210> 4  
 40 <211> 1521  
 <212> ADN  
 <213> Homo sapiens

<400> 4  
 45 gatccctggg ccaaggggaag gagcacatga ggagttgccg aatgtgaaca tgttatctaa 60  
 tcatgagtggt ctttccacgt gctagtttgc tagatgttat ttcttcagcc taaaacaagc  
 120  
 tggggcctca gatgaccttt cccatgtagt tcacagaatt ctgcagtggg cttggaacct  
 180  
 50 gcagccacga aaagatagat tacatatgtt ggagggagtt ggtaattccc aggaactctg  
 240  
 tctctaagca gatgtgagaa gcacctgtga gacgcaatca agctgggcag ctggcttgat  
 300  
 tgccttcctt gcgacctcaa ggaccttaca gtgggtagta tcaggagggg tcaggggctg  
 360  
 55 taaagcacea gcgttagcct cagtggcttc cagcacgatt cctcaacct tctaaccatt  
 420  
 ccaaagggtat tatctttggg ggttgacatt ctttctctgt tttcttttta atctttttt  
 480  
 60 aaaacataga attaatatat tatgagcttt tcagaagatt tttaaaaggc agtcagaaat  
 540  
 cctaactacct aacacaaaaa ttgtttttat ctttgaataa tatgttcttg tttgtccatt  
 600  
 ttccatgcat gcgatgttag gcatacaaaa tacatttttt aaagaatact ttcattgcaa  
 660  
 65 attggaaact tcgttttaaaa aatgctcata ctaaaattgg catttctaac ccataggccc  
 720

acttgtagtt atttaccgaa gcaaaaggac agctttgctt tgtgtgggtc tggtagggtt  
 780  
 cattagaaag gaatgggggc ggtgggaggg ttggtgttct gttctctctg cagactgaat  
 840  
 5 ggagcatcta gagttaagg taggtcaacc ctgacttctg tacttctaaa tttttgtcct  
 900  
 caggtaaatc ctgaccgggt tgttcccccc gacctcgggc accgcctaca tctgggaaa  
 960  
 10 agacattcgc tctgagatga gcaccatccg gcagaacctg ggggtctgtc cccagcataa  
 1020  
 cgtgctgttt gacatgtgag taccagcagc acgttaagaa taggcctttt ctggatgtgt  
 1080  
 gtgtgtcatg ccatcatggg aggagtggga ctttaagcatt ttactttgct gtgtttttgt  
 1140  
 15 tttttctttt tttctttttt atttttttga gatggagtct cgtctgttag ccaggctgga  
 1200  
 ctgtagtggc gcgatctcgg ctactgcaa ccttggcctc ccaggttcaa gcgattctcc  
 1260  
 tgcctcagcc tcccgagtag ctgggactct aggcacacac caccatgccc agctaatttt  
 1320  
 20 tgtgttttta gtagagacgg ggtttcacca tgttgccag gatggtotca atgtcttgac  
 1380  
 ctctgtatcc gccacacctg gtctccaaa gtgctgggaa cacaggcatg agccactgtg  
 1440  
 25 tctggccaca ttttactttc ttgaatatg gcaggctcac ctccgtgaac accttgagac  
 1500  
 ctagtgttc tttgatttta g  
 1521  
 30  
 <210> 5  
 <211> 6519  
 <212> ADN  
 <213> Homo sapiens  
 35  
 <400> 5  
 gaaattgaaa gttgtaactg cctgggtgcat ggtggccagg cctgctggaa acaggttgga 60  
 agcgatctgt cacctttcac ttgatttcc tgagcagctc atgtggttgc tcaactgtgt  
 120  
 40 tctacctga atcttgaaga ttatttttca gaaattgata aagttatttt aaaaagcacg  
 180  
 gggagagaaa aatatgcca ttctcatctg ttctgggcca ggggacactg tattctggg  
 240  
 tatccagtag ggcccagagc ttgacctgcc tccctgtccc caggctgact gtcgaagaac  
 300  
 45 acatctggtt ctatgcccgc ttgaaagggc tctctgagaa gcacgtgaag gcggagatgg  
 360  
 agcagatggc cctggatgtt ggtttgccat caagcaagct gaaaagcaaa acaagccagc  
 420  
 50 tgtcagggtg gccccagagc taccttccct atccctctcc cctcctctc cggctacaca  
 480  
 catgaggagg aaaatcagca ctgccccagg gtcccaggct ggggtcgggt ggtaacagaa  
 540  
 acttgtccct ggctgtgccc ctaggctcctc tgccttcaact cactgtctgg ggctggctct  
 600  
 55 ggagtttgtc ttgctctgtt tttttgtagg tggaatgcag agaaagctat ctgtggcctt  
 660  
 ggcctttgtc ggggatcta aggttgtcat tctggatgaa cccacagctg gtgtggaccc  
 720  
 60 ttactccgc aggggaatat gggagctgct gctgaaatac cgacaagggt cctgatgtgt  
 780  
 atttattctg agtaaagga ctgagagaga gcggggggct tttgagaagt gtggctgtat  
 840  
 ctcatggcta ggcttctgtg aagccatggg atactcttct gttatcacag aagagataaa  
 900  
 65 gggcattgag actgagattc ctgagaggag atgctgtgtc tttattcatc tttttgtccc  
 960

caacatggtg cactaaattt atggttagtt gaaagggtg atgcttaa at gaatggaagc  
 1020  
 ggagaggggc aggaagacga ttgggctctc tggtagaga tctgatgtg tacagtatga  
 1080  
 5 ggagcacagg caggcttga gccaaactctg gctggccctg agacattggg aaagtcacaa  
 1140  
 ctgcctcac cttctttgcc gataataata gtggtgctta cctcatagag gattaaatta  
 1200  
 10 aatgagaatg cacacaaacc acctagcaca atgcctggca tatagcaagt tcccaataa  
 1260  
 aatgctactg ttcttacctc tgtgaggatg tggtagctat atatacaaag ctttgccatt  
 1320  
 ctaggggtca tagccatata gggtgaaagg tggcttccag gtctcttcca gtgcttacc  
 1380  
 15 ctgctaata ctctctagtc cctgtcactg tgacaaatca gaactgagag gcctcacctg  
 1440  
 tcccacatcc ttgtgtttgt gcctggcagg ccgcaccatt attctctcta cacaccacat  
 1500  
 ggatgaagcg gacgtcctgg gggacaggat tgccatcacc tcccatggga agctgtgctg  
 1560  
 20 tgtgggctcc tccctgtttc tgaagaacca gctgggaaca ggctactacc tgaccttgg  
 1620  
 caagaaagat gtggaatcct cctcagttc ctgcagaaac agtagtagca ctgtgtcata  
 1680  
 25 cctgaaaaag gtgagctgca gtcttgggtg tgggctggtg ttgggtctg gcagccagga  
 1740  
 cttgctggct gtgaatgatt tctccatctc caccctttt gccatgttga aaccaccatc  
 1800  
 30 tccctgtctc gttgcccctc tgaaatcata tcatacttaa ggcatggaaa gctaaggggc  
 1860  
 cctctgtctc cattgtgcta gttctgttga atcccgtttt ccttttcta tgaggcacag  
 1920  
 agagtgtg agaggtcct tagaggacat tattatgtca aagaaaagag acttgtcaag  
 1980  
 35 aggtaaagag cttggctaca aatgacctgg tgttctgtct cattactttt caatctcatt  
 2040  
 gaccttaact tttaaactat aaaacagcca atatttatta ggcactgatt tcatgccaga  
 2100  
 gacactctgg gcatgaaaga aagtaatgat aatagtta at tttatatagc gttgttacca  
 2160  
 40 tttacaacct tttttttttt ttttaacctc atcatctcaa ttaaagtga gagagacct  
 2220  
 gggaagaagg taactatatt tattatccca gatgaggga gtgaggcttg tagggaattg  
 2280  
 45 gtagctgatt caaggtcacc cagcaggtaa ataacagtgg tgggaccaga cccaattacc  
 2340  
 aggtatgttt tccctgttac cgcagtacat gcctgagatt tatttgtgtg ttgaagccag  
 2400  
 tggtagctaa tgtatttaca tcccaacctg aaactcctat ccacttattt acctttta  
 2460  
 50 gagctctta actcaagtgc agtctgagga ccagcagcat caggatcact tgggaacttg  
 2520  
 ttagaaatc agcaacctgg gccagctca gacctaccga atcagaatct gtgcatttta  
 2580  
 55 acaaggttct tgagtgttg aacacacatt aaagcatgag aagcattgaa ctagacatgt  
 2640  
 agccaggtaa aggccttggc tgagatggtt ggcaaaggcc tcattgcagc attcattggc  
 2700  
 aggccacagt tcttttggca gctctgttct ctgacctttc accctcagga agcgaggctg  
 2760  
 60 ttcacacggc acacacatgc cagacagggt cctctgaagc cagggtgcc agtgcagtgt  
 2820  
 tcccaggga agctttttcc tttagttctc acacaacaga gcttcttggga agccctcccc  
 2880  
 65 ggcaaagggt ctgggtggctc tgccttctc cgtccctgac ccgttctcac ctcttcttt  
 2940  
 gccatcagga ggacagtgt tctcagagca gttctgatgc tggcctgggc agcgaccatg  
 3000

agagtgcacac gctgaccatc ggtaaggact ctgggggttc ttattcaggt ggtgcctgag  
3060  
cttccccag ctgggcagag tggaggcaga ggaggagagg tgcagaggct ggtggcgctg  
3120  
5 actcaaggtt tgctgctggg ctggggctgg gtggctgcgg gtgtgggagc agcttggtgg  
3180  
cgggttggcc taatgcttgc tggggtgcct ggggctcggg ttgggagcta gcagggcagt  
3240  
gtcccagaga gctgagatga ttggggtttg gggaaatccct taggggagtg gacactgaat  
3300  
10 accagggatg aggagctgag ggccaagcca ggagggtggg atttgagctt agtacataag  
3360  
aagagtgaga gccaggaga tgaggaacag ccttccagat ttttcttggg tagcgtgtgt  
3420  
15 aggaggccag tgtcaccagt agcatatgtg gaacagaagt cttgaccctt gctatctctg  
3480  
cctagtccta atggctggct ttcccagga aggccttctgc ttncatggac ngntagatta  
3540  
accctttatt taggtaaatg agggaacctt ctttataagc ataggaaagg gtgaagaatc  
3600  
20 ttttaagatt cctttactca agttttcttt tgaagaatcc cagagcttag gcaatagaca  
3660  
ccagactttg agcctcagtt atccattcac ccatccacc acccaccac ccactcttc  
3720  
25 atcctcccat cctccattc acccatccac ccatccagct gtccacccat tctacactga  
3780  
gtacctataa tgtgctggc tttggtgata caaaggtgaa taagacatag tcttttctt  
3840  
30 tgccccaac cctcagacca gagatgaaca tgtggaatga cctaaacacc tggaacaggt  
3900  
gtggtgtatg agcggcaggc ctctgatgag aggggtgggg atggccagcc ctcactccga  
3960  
agccctctg agttgattga gccatcttg cattctggc cctgcagatg tctctgctat  
4020  
35 ctccaacctc atcaggaagc atgtgtctga agcccgctg gtggaagaca tagggcatga  
4080  
gctgacctat gtgctgccat atgaagctgc taaggaggga gcctttgtgg aactctttca  
4140  
tgagattgat gaccggctct cagacctggg catttctagt tatggcatct cagagacgac  
4200  
40 cctggaagaa gtaagttaag tggctgactg tcggaatata tagcaaggcc aaatgtccta  
4260  
aggccagacc agtagcctgc attgaggca ggattatcat ggagttagtc attgagtttt  
4320  
45 taggtcatcg acatctgatt aatgttggcc ccagttagcc atttaagatg gtagtgggag  
4380  
atagcaggaa agaagtgttt tctctgtac cacagtacat gcctgagatt tgtgtgtga  
4440  
aaccagtggg acctaacaca ttacatccc aaccttaaac tctatgcac ttatttacc  
4500  
50 tttaatgagc ctctttactt aagtacagtg tgaggaacag cggcatcagg atcacttggg  
4560  
aacttgtag aaatcacga acttgggcc agctcagacc tactgaatca gaatcaggag  
4620  
55 caattctctg gtgtgactgt gtcacagcca ggtatcaact ggattctcat acataggaaa  
4680  
tgacaaacgt ttatggatgg atagtctact tgtgccaggt gctgagattt gttttttgtt  
4740  
60 ttttgatttt ttttaataca ctgtgacctc atttaattct caaaaaaaga tgaaaaaatg  
4800  
aacactcagg aatgctgaca tgagattcag aatcaggggt ttggggcttc aaagtccatc  
4860  
ctctctttat ccatgtaatg cctccctta gagatacaac atcacagacc ttgaaggctg  
4920  
65 aaggggatat aaaagctgtc tggccaagt gtctccaagc ttgacagtgc agcagaatca  
4980  
cctggggata ttattaaaaa taaacatact aagggttggc ttcagggcct gtgaatcaga  
5040

atctctggag gtgaggcctt gaagctctgta tttctattgc atactttgga cacagtggtc  
5100  
tatagactag agtttgaaa tgattgcgct cattcagatt ctcttctgat gtttgaattg  
5160  
5 ctgccatcat atttctagtg ctctatttcc tctgtctcat tctgtcttgg ataaacttate  
5220  
atagtactag cctactcaaa gatttagagc cacagtcctg aaagaagcca cttgactcat  
5280  
tccctgtagg ttccagaataa atttcttctg cgcagtgtct gtcatagctt tttttaaatt  
10 5340  
tttttttatt ttgatgaga ctggagtttt gctcttattg cccaagctgg agtgcagtgg  
5400  
tgcgattttg gctcactgca acctccacct cccaggttca agcgattctc ctgcctcagc  
5460  
15 ctcccaagta gctgagatta caagcatgtg ctaccacgcc cagctaattt tgtattttta  
5520  
gtagagatgg gttttatcca tgttggtcag gctggctctg agctccagac ctcaggtgat  
5580  
ctgcccgcct cggccctcca aagtgcctggg attataggcc tgagccacag cgctcagcca  
20 5640  
taactttaat ttgaaatga ttgtctagct tgatagctct caccactgag gaaatgttct  
5700  
ctggcaaaaa cggcttctct cccaggtaac tctgagaaag tgttattaag aaatgtggct  
5760  
25 tctactttct ctgtcttacg gggctaacat gccactcagt aatataataa tctgtggcagt  
5820  
ggtgactact ctcgtaatgt tgggtgctat aatgttctca tctctctcat ttccagata  
5880  
ttcctcaagg tggccgaaga gagtggggtg gatgctgaga cctcaggtaa ctgccttgag  
30 5940  
ggagaatggc acacttaaga tagtgccttc tgctggcttt ctcagtgcac gagtattgtt  
6000  
cctttccctt tgaattgttc tattgcattc tcattttagt agtgtaggtt tgttgagat  
6060  
35 ggggaaggtt tgttttgttg taaataaaat aaagtatggg attctttcct tgtgccttca  
6120  
gatggtacct tgccagcaag acgaaacagg cgggccttcg gggacaagca gagctgtctt  
6180  
cgcccgttca ctgaagatga tgctgctgat ccaaatgatt ctgacataga cccaggctcg  
40 6240  
ttagggcaag atcaaacagt gtccactagt ttgaatgtga aattctctct catgctctca  
6300  
cctgttttct ttggatggcc tttanccaag gtgatagatc cctacagagt ccaaagagaa  
6360  
45 gtgaggaaat ggttaaagcc acttggtttt tgcagcatcg ngcatgtnat caaacctgan  
6420  
agagcctatc catatcactt tnccttaana gacattaaan atggntcctt aatctctttt  
6480  
gancccatg tatttattat tctttttctg cgggggtcc  
50 6519

<210> 6  
<211> 7378  
55 <212> ADN  
<213> Homo sapiens

<400> 6  
tctagaaaat ttttaggaac agaaaacttt ccagttctct caccctgct caaagagtgt 60  
60 atggctctta cattatatat aactgcctga cttcatacag tatcagtact tagatcattt  
120  
gaaatgtgtc cacgttttac caaaatataa taggggtgaga agctgagatg ctaattgcca  
180  
ttgtgtattc tcaaatatgt caagctacgt acatggcctg ttccatagag tagtctataa  
240  
65 gaaattgatg acttgattca tccgaatggc tggctgtaac acctggttac gcatgaacac  
300

ctcttttcag ttgtctcaag acacctttct tttctgtact tatcagacaa ggactgaaag  
360  
gcagagactg ctactgttag acattttgag tcaagctttt ccttggacat agctttgtca  
420  
5 tgaagccct ttacttctga gaaacttcta gcttcagaca catgccttca agatagttgt  
480  
tgaagacacc agaagaagga gcatggcaat gccgaaaaca cctaagataa taggtgacct  
540  
tcagtgttg cttcttgag aatccagaga gacagacttg ctcagtggga tggatggcaa  
10 600  
agggctctac caggtgaaag gctggaaact tacacagcaa cagtttgttg cccttttgtg  
660  
gaagagactg ctaattgcc aacggagtcg gaaaggattt tttgctcagg tgagacgtgc  
720  
15 tgttttcgcc agagactctg gcttcatggg tgggctgcag gctctgtgac cagtgaaggc  
780  
aggatagcat cctgggtcaag atatggatgc cggagccaga tttatctgta tttcaatccc  
840  
agtctattc cttgccagtt gtgtatccgc tggcaagtta cttctctatg cctcaatctc  
20 900  
ctcatctgta aaatggggat aataatatta cctgcaatac agggttgtta cgaaaataaa  
960  
aatgaatagg tgcttagaat ggggcctgac attagtaagt gcttagtttt gtgtgtgtat  
1020  
25 atgttatttt tattttggag gagaacataa aaaggacaaa gtgtagaaaa actgggtggg  
1080  
tgtattcagc tgtcataaca tgagagttgt tatgccaga tgcacttgac atgtgaattt  
1140  
attagaaaca tgatttttct ctgagttgat gtttaactca aactgataga aaagataggt  
30 1200  
cagaatatag ttggccaaca gagaagactt gttagactat tgtctgcatg tcagtgtttg  
1260  
catgctaact tgcttagtta gaaaggttaa attttttcac tctataaaat caagaaatat  
1320  
35 agagaaaagg tctgcagaga gtctttcatt tgatgatgtg gatattgtta agagcgggag  
1380  
tttgagcat acagagctca agttgaatcc tgactttgct acttattggc tatatgacct  
1440  
tgggcaagct gcttagtctc tctgatccct agttaccttt gtttgttgat gatgaccatt  
40 1500  
gataacacaa ccataaataa tgacaacata gagatagttc tcattatagt agttgttata  
1560  
cagaattact cactcaatgt taattttctg cattgaaatc ccagaacatt agaattgggg  
1620  
45 gcattatttg aatctttaag gttataagga atacatttct cagcaataaa tgggaaggagt  
1680  
tttgggttaa cttataaagt ataccctagt catttttttt tcagagaaga tatggtagaa  
1740  
50 agtcttagga ggttgaagaa ggaattggat atttattctt tctgagacta tcatgggaga  
1800  
taatgactat ggttgtccat gattggagcc gttgctgtag agttggtttt attatagttg  
1860  
aggatttgaa tgggccatgt gttctcagac ctcagattaa aatgagaaaa ctgaggccag  
1920  
55 tggggagcgt gacttcacat gggtagactt gtgctagaga cagaaccagg attcaggact  
1980  
tctggctcct ggtcctgggt tcatggccca atgtagtctt tctcagtctt caggaggagg  
2040  
aagggcagga ccagtggtc tgagtcaccc tgaatgtgag cactatttac ttcgtgaact  
60 2100  
tcttggctta gtgcctctgc cagggtggca taacctctgg ccttgtgttg ccagagaaaa  
2160  
ggtttagttt tcaggctcca ttgcttccca gctgccaaga atgccttggg gcagcacagt  
2220  
65 catagccct gcattcctca ttgccgtgct ggttggctcg ggaggtgggc tggactcgta  
2280  
gggatttgcc ccttggcctt gtttctaaca cttgccgttt cctgctgtcc ccctgcccc  
2340

tccactgcct gggtaaagat tgtcttgcca gctgtgtttg tctgcattgc ccttgtgttc  
2400  
agcctgatcg tgccaccctt tggcaagtac cccagcctgg aacttcagcc ctggatgtac  
2460  
5 aacgaacagt acacatttgt caggatatgt tgtcttctac atcccaggag ggggtaagat  
2520  
tcgagcagac caaagatgtt tacgagggcc aagggaatgg acttcagaat tacacggtgg  
2580  
aatgaatttt actgctgcgg ctccaggctcc tgtataagct aatactgcat gcatagaaca  
2640  
10 gcagcgaact aaccctgaat aataggccag tcttctgttg agcctttcag cctctctcct  
2700  
cttcactcta ctgttgctag gaacagccac atgtgtttta ggtgaaataa tccacccttg  
2760  
15 caaaaatcca tgattaagtt ataaaatatt tggatttgg gagctgtgtt ttaattctgt  
2820  
aactgagtca cagggcacac tgtcaaagca tagaacctcc agagacttgt tttctgcaaa  
2880  
gtataattca tgtaattatt atctattctg ttatatattgg gatgttaggt agtgtttgtt  
2940  
20 ctttagataa aaatatcccc cactctgtaa caatacatta aatcaaagaa aaggacaaag  
3000  
gatttttctg ggtcttgta gcaggagctt tcttcagctc tgaaagattt gtagacctgt  
3060  
25 agatggggga actgtgtcag tgatacaaaa gggaagcatt taaaaaaaaa aagtatatat  
3120  
atatatatat atatattgaa tgtgaattgg cctctttttc tctaagccca cattttcttc  
3180  
ttacatagtt caggtttact ttattttttc ctttccggct gctgacctg tattgcccg  
3240  
30 agttgtggaa catagcatgt gtttgtgacc tgtgcctgtt atttttgtgc tttctagtgt  
3300  
tgcattgcaa gagtacaag ttttcttgcc ctttcttgga aaatcctgct tgtctgtgcc  
3360  
35 aaaggataa ttgtgaaagc acttttgaaa tacttaataa gttgattttc ttcaaatata  
3420  
aaaaaatata taaatgtatc tgtgtatgta catgtgtgta cacatacaca cttttatata  
3480  
tacagcccat ttaaaacaag ctccactttg gagtgctcta cgtcaccctg atgccgaata  
3540  
40 cagggccaga gtctgagatc cttctgggtg gtttctgtgt tttgttcatt tctgttttaa  
3600  
gagcctgtca cagagaaatg cttectaaaa tgtttaattt ataaaaacat ttttatctct  
3660  
45 cgattactgg ttttaataa ttactaagct ggctgcctct catgtacca cagcaatgat  
3720  
gctcctgagg acacgggaac cctggaactc ttaaagccc tcaccaaaga ccctggcttc  
3780  
gggacccgct gtatggaagg aaacccaatc ccgtgagtgc cactttagcc ataagcaggc  
3840  
50 ttcttgtgct tgttgcttg tttgatttct aatatgctgc atttatcaac tgcattgccac  
3900  
attgtgaccg ccagcatttg ccctttgaat tattattatg ttttatttac aaaaagcgaa  
3960  
55 ggtagtaacc gaactaaatt atctaggaac aaacgtttgg agagtcttct aacaccgtgc  
4020  
aaagcacgtc attacagaca tttgtttact gatttagaac cttaatatat aatttaaata  
4080  
gcactttaca ctactgatg aaatgctttt ctttcttttc tctcccagcc cctgtactta  
4140  
60 agtgcttcaa taggctctca ttatatatga tttttagggt ttgcttatca gcttcttcgc  
4200  
ttttataatc tgaaaagatg gcatatgaat ttttataaaa agggacactt tcttcttctc  
4260  
65 aaattgtata tttttattgt actttccttc aaaacccctt tttaaaaagt aagcagtggg  
4320  
taaataaatt cagtgaagca tccatatgac ccttaagtga gtgtagggga agggagggtca  
4380

ccagatcact gtgagtgaag atggtggaga ggtgaggatc ttatgaggcc gtgctcaagg  
 4440  
 ctggtagagg tgggttagtg tttccagggt taggcagaat ctgagctgag gtcattgaaac  
 4500  
 5 aacagtgtac tctgaaaaat tatggcaagg tgggaagggtg ctggagaatt ggagaggggg  
 4560  
 caaacttgac tttcaagttt caatgggaag ataggtgact ctgcacacca cagaacagtg  
 4620  
 10 agcatgataa cctgtttata caaggttcta gagcagattt ctaaattggat agctactgtg  
 4680  
 tgcttgtttg ttcttaatta gtattggata gttactaaat acttgtagt acttagtaca  
 4740  
 taatgggttg taaatcctag cagctaatat tggttcccaa ataaccagat gacaaggata  
 4800  
 15 gagaaggaca cagacacggc ctatctggat ttcattgtgc ctttcatttt ccacatgaag  
 4860  
 gttgtgtagg gaagatagaa gcatgagatg agatgataat atagttatct ggattcatca  
 4920  
 ctggccagct gaaccatag aactcatgga ttgatgctag cttaggaagg ctctgttaga  
 4980  
 20 gccagaactg ggctgagagc cagcccatag agacaaaaga ggcccgcccc tgacatcaga  
 5040  
 gggttcaaac atgatgtctg agccccacct acagtctgcc ggaggtggtt ggaaggaga  
 5100  
 25 gcccttatcc ttacaattct tactgaaatt caaattttta ggttttgcaa aaaaatggtg  
 5160  
 gacctgaagg aaatttgaca ggagcatgtc tcagctgtat ttaattttgt ctgaccaat  
 5220  
 ccccttttga atgttcagag tgtaagcttc aggagggcag cgcgtcttag tgtgactttt  
 5280  
 30 ctggtcagtt caggtgcttt aaggagacaa ttagagatca atctggaaaa cttcatttga  
 5340  
 atttttaata cataagaaaa caataagaaa tagttaaaaa tatatatata taatatatat  
 5400  
 35 atgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtatatatat atattttatt tattttttt  
 5460  
 tttttgagat ggagtctcgc tctgttgccc aggcctggagt gcagtggctc aatcttggtc  
 5520  
 cactgccacc tctgcctccc aggttcaagt gattctccta cctcagcctc ctgagttagt  
 5580  
 40 gggattacaa gcatgtgcca ccacactggc taatttttct aatttttagta gagatggagt  
 5640  
 ttcaccatgt tggacaggat ggtcttgaac tcctgactta gtgatccacc cgccttcgcc  
 5700  
 45 tccc aaagt ctgggattac aggcagagc catcgtgcct ggcaattata tttaatatt  
 5760  
 aataataagg aaataattgc tgtaacttta ctttaaatg tggaattctg aaactggaag  
 5820  
 50 ggaactggaa atgacttggt gaatcaaatc attttaaact tttattttgc cagtggaaaa  
 5880  
 aataagcccc caaaagagca ggggacctgc tgatgtccca cagtaattca gagctggaga  
 5940  
 tgagggtgaa ggcttttgtgt ctatctcca gggaaaattt gtagacagcg tagctcttta  
 6000  
 55 tgtgacgagc attctcacc cagtcacccc ccaattctct actcatttga gaacataaat  
 6060  
 tggatcttgc cagtctctac tcatttttca gcacatcgag cataagatcc agactctttc  
 6120  
 ccaggcctct ctcatctggc tctctctctc ctcctttatc attactcttc ttcgtagctt  
 6180  
 60 atcctactcc agccatgctg tcttctatt attcctaaaa agtagaaatg catttcttc  
 6240  
 tagggccttt gtacctgcac ttgccatcgc ttttgcacag aatgttcttt ttgccaagct  
 6300  
 65 tttgcccage ttgttctcca tcattgttat gttttggctg aaatgtcttc tcttagtagg  
 6360  
 ttcattctcc ccagtcactg tctttttatt ttgctttatt ttgggccatc taaggttatc  
 6420



ttattagtgt atttgttgtt cgtctcctcc atgggcatac acctccatga aggcagggtat  
 6480  
 tttcacctta ggccctcgaata tctactggac agcatctggc acgtagtaga tgctcaacga  
 6540  
 5 atgtttgttg tgtgagcaaa tgggttggtt atttgattga actgagttca gtatgtaaa  
 6600  
 atttagggcc tctttgcatt ctattttact tatgtataaa atgatacata atgatgatat  
 6660  
 10 aaatgatgtc acagtgtaca aggctgttgt gggatcaagc aatcaaata gaatcatgctt  
 6720  
 gtcttttcca aatggtgagg gaatagatgc atgtttgttg ttgttacgga atgatcctgt  
 6780  
 gctcctgagg caacagaaag gccaggccat ctctggtaat cctactcttg ctgtcttccc  
 6840  
 15 tttgcagaga cagccctgc caggcagggg aggaagagtg gacactgcc cagttcccca  
 6900  
 gaccatcatg gacctcttcc agaattggaa ctggacaatg cagaaccctt cactgcatg  
 6960  
 20 ccagtgtagc agcgacaaaa tcaagaagat gctgcctgtg tgtccccag ggcaggggg  
 7020  
 gctgcctcct ccacaagtga gtcactttca gggggtgatt ggcagaaag ggtgcaggat  
 7080  
 gggctggtag ctctcgcttg gaagcaggaa tgagttagat atcatgttg gagggtctgt  
 7140  
 25 ttcagtcttt tttgtttttt gttttttttt ctgaggcgga gtcttgcctt gtcgccagg  
 7200  
 ctggagtgtc gtggcatgat ctgcctcac tgcaacctcc acctcccagg ttcaagcat  
 7260  
 tctcctgctt cagcctcctg agtagctggg attacaggca cgcaccacca tgtctggcta  
 7320  
 30 atttttgtgt ttttagtaga gatagggtt cgcctgttg gctaggctgg tctggaat  
 7378  
  
 35 <210> 7  
 <211> 5689  
 <212> ADN  
 <213> Homo sapiens  
  
 40 <400> 7  
 gaattcctga cctcaggtaga tccaccgccc tggcctccc aaagtgttg gattacaggc 60  
 gtgagccact agccccagcc ctgtttcagt ctttaactcg cttcttgtca taagaaaaag  
 120  
 catgtgagtt ttgaggggag aaggtttga ccacactgtg cccatgcctg tcccacagca  
 180  
 45 gtaaaagtcac aggacagact gtggcaggcc tggcttccaa tcttggtctt gcaacaaatg  
 240  
 agctggtagc ctttgacagg cctgggcttg tttcttcacc tctgaattag ggaggctgga  
 300  
 50 ccagaaaact cctgtggatc ttgtcaactc tggatttctt agagactctg tttgggaagg  
 360  
 agtctgagc catttttttt ttcttgagaa ttccagggaag aggagtgtt atgatagctc  
 420  
 tctgctgctt ttatcagcaa ccaaattgca ggatgaggac aagcaattct aatgagtagc  
 480  
 55 aggaactaaa agaaggcttg gttaccactc ttgaaaataa tagctagtcc aggtgcgggg  
 540  
 tggctcacac ctgtaatctc agtatttttg gatgccagg tggactgatc acctaaaggtc  
 600  
 60 aggaagttcg aaccagcttg gccaatgttg cgaaccctg tctctactaa aaattcaaaa  
 660  
 attagccagg catggtggca catgcctgta atcccagtt cttgggaggc tgaagcagg  
 720  
 gaattgcttg aacctgggag gtggaggtcg caggagacca aaattgcgcc actgtactcc  
 780  
 65 agcctgagca acacagcaaa actccatata aaaaaataa atgaataaaa taacagctaa  
 840

tctagtcatc agtataactc cagtgaacag aagatattatt aggcatagtg aatgatgggtg  
900  
cttcctaaaa atctcttgac tacaaagaat ctcatattcaa tgtttattgt ttagatgttc  
960  
5 agaataaatt cttgggaaag accttggcctt ggtgtaagtg aattaccagt gccgagggca  
1020  
gggtgaacca agtctcagtg ctggttgact gagggcagtg tctgggacct gtagtcaggt  
1080  
ttccggtcac actgtggaca tggtcactgt tgccttgat ttgttttctg tttcaattct  
1140  
10 tgtctataaa gacccttatg cttggtttct atgtgatgac agagaaaaca aaacactgca  
1200  
gatatccttc aggacctgac aggaagaaac atttcggatt atctggtgaa gacgtatgtg  
1260  
15 cagatcatag ccaaaagggtg actttttact aaacttggcc cctgccgtat tattactaat  
1320  
tagaggaatt aaagacctac aaataacaga ctgaaacagt gggggaaatg ccagattatg  
1380  
gcctgattct gtctattgga agtttaggat attatcccaa actagaaaag atgacgagag  
1440  
20 ggactgtgaa cattcagttg tcagcttcaa ggctgaggca gcctggctca gaatgaaaat  
1500  
agaaatggat tcaacgtcaa attttgccac ttagtagcaa cttgaccagg taactggtta  
1560  
25 tcctttttaa gccttagttt atctaaattg tgatattaat gttgctctta taagtttgtc  
1620  
atgaggacta aattaaatgg tgtacataga gtgccttggg tactctctga tgggggactc  
1680  
catgataatt tgtggtctca tggagggagc tctgggaagg tttaggagcc tgccttggct  
1740  
30 ctgcagcctt gggagagcct tctagcttcc caggacatgg cagcctagtg ttgaatgctt  
1800  
ggctcagcaa atgtttgttc tcgtttcctt cccatcaact tggtcagttg gggcttttca  
1860  
35 gttaggagta tctcagtgc tttaaatggc atgggcatgc tggagtata gtgaccatga  
1920  
gtttctaaaga aagaagcata atttctccat atgtcatcca caattgaaat attattgtta  
1980  
attgaaaaag cttctaggcc aggcacggtg gctcatgcct gtaatcccag cactttagga  
2040  
40 ggccaaggcg ggtggatcac ttgaggtcag gagtttgaga ccagcctggc caacatgggg  
2100  
aaaccctgtc tctactaaaa atacaaaata agctgggctt ggtggtgctg gcctgtaatc  
2160  
45 ccagctactt gggaggctga ggcaggagaa ttgcttgaat ctgggaggcg gaggttgagc  
2220  
tgagctgagt tcatgccatt gcattccagc ctgggcaaca agagcgaaac catctcccaa  
2280  
aagaaaaaaa aaagaaagaa aaagcttcta gtttggttac atcttggctc ataagggtgt  
2340  
50 ttgtaaatgt gtttaaccac aggcctggtt ctcatataag taatagggtta tttatgatgg  
2400  
agagaaggct ggaagaggcc tgaacacagg cttcttttct ctgacacaac cctacaaggc  
2460  
55 cagctgattc tagggttatt tctgtccgtt ccttatatcc tcaggtggat attactcct  
2520  
tttgcacat taggaatagg ctcaagtctt tctttgaact gattttttgt tcttttgtct  
2580  
ctgcagctta aagaacaaga tctgggtgaa tgagtttagg taagttgctg tctttctggc  
2640  
60 acgttttagct cagggggagg atggtgttgt aggtgtcttg gattgaagaa agccttgggg  
2700  
attgtttgtc actcacacac ttgtgggtgc catctcactg tgaggaggac agaagccctg  
2760  
65 tgaacatgtg gagcacacag gggcacagac agatttagat taggcctgct ttatagagtt  
2820  
tctgcctaga gcatcatggc tcagtgccca gcagcccctc cagaggcctc tgaaatattt  
2880

gatatactga ttctcttgag gagaatcaga aatctcctgc aggtgtctag ggatttcaag  
 2940  
 taagtagtgt tgtgaggga atacctactt gtactttccc cccaaaccag attcccagg  
 3000  
 5 cttcttaagg actcaaggac aatttctagg catttagcac gggactaaaa aggtcttaga  
 3060  
 ggaaataaga agcgccaaaa ccactctctt gcactgtatt tcaaccatt tgccttctg  
 3120  
 10 ggttttgaag gaacagggtg gactggggac agaagagttc ttgaagccag ttgtccatc  
 3180  
 atggaaaatg agataggtga tgtggctacg tcagggggcc cgaaggctcc ttgttactga  
 3240  
 ttccgtctt ttctctctgc cttttcccca agggccaggc cccctggatc tctgggcaga  
 3300  
 15 gcagacgcag gccctataa tagccctcat gctagaaagg agccggagcc tgtgtataag  
 3360  
 gccagcgcag cctactctgg acagtgcagg gttccactc tcccaactcc ccactctgctt  
 3420  
 20 gcctccagac ccacattcac acacgagcca ctgggttggg ggagcatctg tgagatgaaa  
 3480  
 caccattctt tcctcaatgt ctgagctatc taactgtgtg tgtaatcagg ccaggctctc  
 3540  
 cctgctgggc agaaaccatg ggagttaaga gattgccaac atttattaga ggaagctgac  
 3600  
 25 gtgtaacttc tctgaggcaa aatttagccc tcctttgaac aggaatttga ctgagtgaac  
 3660  
 cttgtacaca ctgcactga gtctgctgct gatgatactg tgcacccac tgtctgggtt  
 3720  
 ttaatgtcag gctgttcttt taggtatggc ggcttttccc tgggtgtcag taatactcaa  
 3780  
 30 gcacttctc cgagtcaaga agttaatgat gccatcaaac aaatgaagaa acacctaaa  
 3840  
 ctggccaagg taaaatatct atcgtaagat gtatcagaaa aatgggcatg tagctgctgg  
 3900  
 35 gatataggag tagttggcag gttaaacgga tcacctggca gctcattgtt ctgaatatgt  
 3960  
 tggcatacag agccgtcttt ggcatttagc gatttgagcc agacaaaact gaattactta  
 4020  
 40 gttgtacgtt taaaagtga ggtcaaaaac aaatccagag gccaggagct gtggctcatg  
 4080  
 cctgtaatcc tagcactttg ggaggccgaa gcgggtggat cacttgaggt caggagtctg  
 4140  
 agaccagcct ggcctacatg acaaaacccc gtatctacta aaaatacaaa aaaattagct  
 4200  
 45 gggcttgggtg gcacacacct gtaatccag ctacttggga ggctgaggca ggagaattgc  
 4260  
 ttgaaccctg taggaagagg ttgtagttag ccaagatcgc accgttgac tccagcctgg  
 4320  
 gcaacaagag caaaactcca tctcaaaaaa caaattaaat ccagagattt aaaagctctc  
 4380  
 50 agaggctggg cggggtggct tacacctgtt atcccagcat tttgggatgc cgaggcgggc  
 4440  
 aaagcacaag gtcaggagtt tgagaccagc ctggccaaca tagtgaaacc ctgtctctgc  
 4500  
 55 taaaaacata gaaaaattag ccgggcatgg tggcgtgcgc ctgtaatccc agctactcgg  
 4560  
 gaggtgagg tgagagaatt acttgaaccc gggaggcgga gggtgcagt agcccagatt  
 4620  
 gcaccaactgc actccagcct gggcgacaga gcaagactcc atctcaaaaa aagctctcag  
 4680  
 60 aacaaccagg tttaaaaatt tggctcagtt gtaataaac tgggtttcaa acatactttg  
 4740  
 ctgaaacaat cactgactaa ataggaaatg aatctttttt tttttttttt aagctggcaa  
 4800  
 65 gctggctgtg aggacctgat aagtactcac ttcatctctc tgtgtctcag gtttccatt  
 4860  
 tttaggtgag aattaagggg ctctgataaa acagacccta ggattgtgga cagcagtgat  
 4920

agtcctagag tccacaagtc tgcttttgag tgatgggccc atgtatctgg cacatctgca  
 4980  
 ggcagagcgt ggttctggct cttcagatga tgccgggtgga gcactttgag gagtcctcac  
 5040  
 5 cccaccgtga taaccagaca ttaaaatctt ggggctttgc atcccaggat ttctctgtga  
 5100  
 ttcttcttag acttgtggca tcatggcagc atcactgctg tagatttcta gtcacttggc  
 5160  
 tctcaggagc cgtttattta atggccttcac atttaatttc agtgaacaag gtagtggcat  
 10 5220  
 tgctcttcac agggccgtcc tgttgccac aggttcaga ttgactgttg ccccttatct  
 5280  
 atgtgaacag tcacaactga ggcaggtttc tgttgtttac aggacagttc tgcagatcga  
 5340  
 15 tttctcaaca gcttggaag atttatgaca ggactggaca ccaaaaataa tgtcaaggta  
 5400  
 aaccgctgtc tttgttctag tagctttttg atgaacaata atccttatgt ttcctggagt  
 5460  
 actttcaact catggttaaag ttggcagggg cattcacaac agaaaagagc aaactattaa  
 20 5520  
 ctttaccagt gaggcagtac ggtgtagtgt agtgattcag agaatttgct ttgccaccag  
 5580  
 acataccagg taaccttgac taagttactt aacctatcta aacctcagtt ccctcatctg  
 5640  
 25 tgaaatggag acagtaatca tagctatttc caaactgttg tgagaattc  
 5689  
  
 <210> 8  
 30 <211> 645  
 <212> ADN  
 <213> Homo sapiens  
  
 <400> 8  
 35 gaattcaatg agttaaaggt ataaggtcct caccacagcg cctgccaca tagtcagtga 60  
 tcactatgtc ctgaacactg taattacttc gccatattct ctgatcatag tgttttgct  
 120  
 tggtagtgta ctagaatttc ttctgagggt ttatgggcat ggttgggtggg tatgcacctg  
 180  
 40 cctgcaggag cccggttttg gggcattacc ttgtacctgg tatgttttct ttcagggtgtg  
 240  
 gttcaataac aagggtggc atgcaatcag ctctttcctg aatgtcatca acaatgccat  
 300  
 tctccgggcc aacctgcaa agggagagaa ccctagccat tatggaatta ctgctttcaa  
 45 360  
 tcatccctg aatctacca agcagcagct ctgagagggt gctctgtaag tgtggctgtg  
 420  
 tctgtataga tggagtggg caaggagag ggttatggag aaggggagaa aaatgtgaat  
 480  
 50 ctcatgttag gggaacagct gcagagaccg ttatattatg ataaatctgg attgatccag  
 540  
 gctctgggca gaagtgataa gtttacgaat tggctggttg ggcttctga actgcagaag  
 600  
 agaaaatgac actgatatgt aaaaatcgta acatttagtg aattc  
 55 645  
  
 <210> 9  
 60 <211> 1664  
 <212> ADN  
 <213> Homo sapiens  
  
 <400> 9  
 65 gaattcatat aaagtgaagt caaaaattgt taattaaatt ataatttaata tataagtgtt 60  
 taatcagttt gatttgttta aaaaccactg ttttaattt ggtggaatat gttttatta  
 120  
 gcttgtatct ttaattccta aattaagctg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg  
 180

240 tgtgtgtgtg aagtttaaag ccaggatgag ctagttaaag gtatgcagcc tttggagtca  
 300 tacagatctg gggttgaatc tgggtctctaa actttataga tgtatgatat taaatgagggc  
 5 agttcatgta aattgccaaag cccagcactc agcacagagt tgatatttca cacacattag  
 360 atacccttcc tgtatgtgga gcatggcagt tcctgtttct gctttactcc tacaggatac  
 420 taatatagga cactaggatc tttataccaa gaccccatgt aatgggctta tgagaccatt  
 10 480 cttcttataa aaatctgaca gaatttttgt atgtgttaga tcaataggct gcatactgtt  
 540 attttcaagt tgatttacag ccagaaatat taatttattt gagtagttac agagtaatat  
 600 ttctgtctctc atttagtttt caagccccac tagtcctttg tgtgtgaaaa tttacaactt  
 15 660 actgctctta caaggctatg aacagtggac caaagtgaat gccattaacc actctgactt  
 720 ccttcattag ttttattgtg acagtggact cttttgacct cagtaatacc agtttggcat  
 20 780 ttacattgtc atatttttag acttaaaaat gatcatctta accctgaata aaatgtgtct  
 840 ggtgaacaga tgttttttct tgggctgtgc ctcagatate tcctgtgtgt gtgtacgtgt  
 900 gtgtttgtct gtgtgtccat gtctcactg attgagccct agctgcatca aaagaccctt  
 25 960 cagattttca cacgcttttt ctctccagga tgaccacatc agtggatgtc cttgtgtcca  
 1020 tctgtgtcat ctttgcaatg tccttcgtcc cagccagctt tgtcgtattc ctgatccagg  
 30 1080 agcgggtcag caaagcaaaa cacctgcagt tcatcagtgg agtgaagcct gtcactact  
 1140 ggctctctaa ttttgtctgg gatatggtaa ggacacaggc ctgctgtatc tttctgatgt  
 1200 ctgtcagggc catggattga tatggataag aaagaaagag ctctggctat catcaggaaa  
 35 1260 tgttccagct actctaaaga tgtatgaaaa agaaatagcc agaggcaggt gatcactttc  
 1320 atgacaccaa acacagcatt gggataccaga gttcatgtca caccagaggg aaaattctgt  
 40 1380 acacaatgat gaaaattaat accactacca cttaagttcc tatgtgacaa ctttcccaag  
 1440 aatcagagag atacaagtca aaactccaag tcaatgcctc taacttctct gatgggtttt  
 1500 aacctccaga gtcagaatgt tctttgcctt actaggaaaag ccactctgtca tttgaaaact  
 45 1560 ctgtacattt tatcagcagc ttatccatcc attgcaaata tgtttttgtg ccagccacaa  
 1620 tatattgctt ctatttggac caataggggg atttgaagga attc  
 50 1664

&lt;210&gt; 10

&lt;211&gt; 1279

55 &lt;212&gt; ADN

&lt;213&gt; Homo sapiens

&lt;400&gt; 10

60 gaattctcat aattgtccta tcgtcaagtc tttatttctg cattttactg cttgatacac  
 120 tgtcaggaca gactttaaaa ttattctcag tgcgatgaaa caattctgac attcatgtta  
 180 tgagcagtta cctcataaat agattacatg tgagattgaa cttgggcaga ctataatata  
 240 gcattaatga cgaaacagac acagtcatct tcgggaagaa gaatagaggc ttatttctgt  
 65 300 cctgtgaaat taaaattact ctgactggga atccatcgtt cagtaagttt actgagtgtg

acaccttggc ttgactgttg gaaagacaga aagggcattgt agtttataaa atcagccaag  
 360  
 gggaaaatgc ttgtcaaaat gtattgtcgg gtattttgat taatagttaa tgtggcttca  
 420  
 5 ttaattcaga gttactctcc aatatgttta tctgcccttt cttgtctgat aatggtgaaa  
 480  
 acttgtgtga tgcattgtat atttgattta ggggtgaact ggatgtcttt gttttcactt  
 540  
 ttagtgcaat tacgttgtcc ctgccacact ggtcattatc atcttcatct gcttccagca  
 600  
 10 gaagtcctat gtgtctctca ccaatctgcc tgtgctagcc cttctacttt tgcgtatgg  
 660  
 gtaagtcacc tctgagtga gggagctcac agtggataag gcatttgggtg cccagtgtca  
 720  
 15 gaaggagggc agggactctc agtagacact tatctttttg tgtctcaaca ggtggtcaat  
 780  
 cacacctctc atgtaccacg cctcctttgt gttcaagatc cccagcacag cctatgtggg  
 840  
 gctcaccagc gtgaacctct tcattggcat taatggcagc gtggccacct ttgtgctgga  
 900  
 20 gctgttcacc gacaatgtga gtcattgcaga gagaacactc ctgctgggat gagcatctct  
 960  
 gggagccaga ggacagtgtt taattgtgat cttattccac ttgtcagtgg tattgacact  
 1020  
 25 gctgactgcc ttgtcctgtc ttcagagtct gtcttccctg agaaggcaaa gcacctttct  
 1080  
 ttcttctgtt gccttacatt ttgctggta agcctttcag tttcttttga cagtttttt  
 1140  
 tacttcttct ttttttcaat gttgctctta ccaagagtag ctctctctgcc ttccacttta  
 1200  
 30 cacatgagag ctggggcagc ccattcagtc ctaaggcttt taccatcacc tctcttgggtg  
 1260  
 tttttattgt catctctaa  
 1279  
 35  
 <210> 11  
 <211> 1124  
 <212> ADN  
 40 <213> Homo sapiens.  
 <400> 11  
 ttaattgat tcaactaggat atatgctact gaaaggggaa tctgcttaaa gtgctttctg 60  
 45 atatttatta ttactaaaac ttagaattta ttaaaaaaac tgactgtgaa aaattacttg  
 120  
 ggtcgtttgc ctttttaaaa ggatttttgg catgtctcat taaaaaaaga aatactagat  
 180  
 atcttcagtg aagttacaaa tcgaatacac attggctctg aaattctgat tgatactggg  
 240  
 50 tcataaaaaag ttttcccaaa tcagacttgg aaagtgatca ctctcttgtt actctttttt  
 300  
 ccttgtcatg ggtgatagcc atttgtgttt attggaagat cgggtgaattt taaggaacat  
 360  
 aggcccaaat ttgaggaagg gccatgggtt ttgatccctc cattctgacc ggatctctgc  
 420  
 55 atttgttcta ctaggggaga atcgctttgt gtcaccatta tcttgggact tgggtgggacg  
 480  
 aaacctcttc gccatggccg tgggaaggggt ggtgttcttc ctcttactg ttctgatcca  
 540  
 60 gtacagattc ttcatcaggc ccaggtgagc tttttcttag aaccctgga gcacctggtt  
 600  
 gagggtcaca gagggggcgc acagggaaac actcaccaat ggggggttga ttgaactgaa  
 660  
 ctcaaaaatat gtgataaaac tgatttttct gatgtgggca tcccgagcc ccctccctgc  
 720  
 65 ccattcctgga gactgtggca agtaggtttt ataatactac gttagagact gaatctttgt  
 780

cctgaaaaat agtttgaaag gttcattttt cttgtttttt cccccaagac ctgtaaatgc  
 840  
 aaagctatct cctctgaatg atgaagatga agatgtgagg cgggaaagac agagaattct  
 900  
 5 tgatggtgga ggccagaatg acatcttaga aatcaaggag ttgacgaagg tgagagagta  
 960  
 caggttacaa tagctcatct tcagtttttt tcagctttat gtgctgtaac ccagcagttt  
 1020  
 gctgacttgc ttaataaaaag ggcatgtgtt cccaaaatgt acatctatac caaggttctg  
 1080  
 10 tcaattttat tttaaaaaca ccatggagac ttcttaaaga attc  
 1124

15 <210> 12  
 <211> 729  
 <212> ADN  
 <213> Homo sapiens

20 <400> 12  
 gaattcccat tctcgaatac attgggttta tatgcttaca tttatgtgtt agttattaaa 60  
 acatactaatt attgtatatac tagtcaaaac tgaggtagag agaataaatg gttgattttg  
 120  
 agttttgagtt tcatagtcca aaaagctgat atattgcctg tgttcaagag ggtctatatc  
 180  
 25 agccctctag atgccagcat ctccaaattt tacttttttg gaatctgtac agtatttgca  
 240  
 atatttttat tacaaatttc tactctgtgg aatttaattt ttaaaatacc tgcaatacat  
 300  
 30 atatatgttg aatagatgaa aaattatgta gataataatg aatgatcagg ttctaaaaag  
 360  
 acagggttaa aagtaagttc actttttatt tgagcttcag aatcattcag aagccagtcg  
 420  
 ccacaaacgc agaccaaggc tcttggcaca tcaaatatgc ctatggctta gggttattga  
 480  
 35 caagtcttat gttgcagtgt atgtgggtta tagtcctgcc ttccacagtt gcttgggaga  
 540  
 gctgtgagtc actgaggctt atgaatgttt acattttgtt tgttgcagat atatagaagg  
 600  
 40 aagcgggaagc ctgctgttga caggatttgc gtgggcattc ctctgggtga ggtaaagaca  
 660  
 ctttgtctat attgcgtttg tccctattag ttccagactat ctctacccaa tcaagcaacg  
 720  
 atgctcgtt  
 45 729

<210> 13  
 <211> 731  
 50 <212> ADN  
 <213> Homo sapiens

<400> 13  
 acatgtgcca gtactgtgga gaggcgaagc tttggagtca aacacaaatg ggtttgcac 60  
 55 ctggccctac caattatgag ctctgagcca tgggcaagtg actaactccc tgggcctcag  
 120  
 tttctctgta acatctgtca gacttcatgg gtccagggtga ggattaaagg agatcatgta  
 180  
 tttacagcac atggcatggt gcttcacata aaataagtat ttagtaaatg ataactggtt  
 240  
 60 ccttctctca gaaacttatt tctgggcctg ccaggggccg ccttttttca tggcacaagt  
 300  
 tgggttccca gggttcagta ttcttttaaa tagttttctg gagatcctcc atttgggtat  
 360  
 65 tttttctgc tttcagggtt ggagatggtt atacaatagt tgtacgaata gcagggtcca  
 420  
 acccggaact gaagcctgtc caggatttct ttggacttgc atttcctgga agtgttctaa  
 480

```

aagagaaaca ccggaacatg ctacaatacc agcttccatc ttcattatct tctctggcca
540
ggatattcag catcctctcc cagagcaaaa agcgactcca catagaagac tactctgttt
600
5 ctcagacaac acttgaccaa gtaagctttg agtgtcaaaa cagatttact tctcagggtg
660
tggattcctg ccccgacact cccgcccata ggtccaagag cagtttgtat cttgaattgg
720
tgcttgaatt c
10 731

<210> 14
<211> 3501
15 <212> ADN
    <213> Homo sapiens

<400> 14
gaattcttca acagggaaaa cagctagctt gaaaacttgc tgaaaaacac aacttgtgtt 60
20 tatggcattt agtaccttca aataattggc ttgcagata ttggataccc cattaaatct
120
gacagtctca aatttttcat ctcttcaatc actagtcaag aaaaatataa aaacaacaaa
180
25 tacttcata tggagcattt ttcagagttt tctaaccag tcttattttt ctagtcagta
240
aacatttgta aaaatactgt ttcactaata cttactgtta actgtcttga gagaaaagaa
300
aaatatgaga gaactattgt ttggggaagt tcaagtgatc tttcaatata attactaact
360
30 tcttccactt ttcccaaaat ttgaatatta acgctaaagg tgtaagactt cagatttcaa
420
attaatcttt ctatatcttt taaatttaca gaatattata taaccactg ctgaaaaaga
480
aaaaaatgat tgttttagaa gttaaagtca atattgattt taaatataag taatgaaggc
35 540
atatttccaa taactagtga tatggcatcg ttgcatttta cagtatcttc aaaaatacag
600
aatttataga ataatttctc ctcatttaat atttttcaaa atcaaagtta tggtttccctc
660
40 attttactaa aatcgatttc taattcttca ttatagtaaa tctatgagca actccttact
720
tcggttccctc tgatttcaag gccatatttt aaaaaatcaa aaggcactgt gaactatttt
780
gaagaaaaca caacatttta atacagattg aaaggacctc ttctgaagct agaacaatc
45 840
tatagttata catcttcatt aatactgtgt taccttttaa aatagtaatt ttttacattt
900
tcctgtgtaa acctaatgtt ggtagaaatt tttaaccaact ctatactcaa tcaagcaaaa
960
50 tttctgtata ttccctgtgg aatgtacctc tgtgagtttc agaaattctc aaaatacgtg
1020
ttcaaaaatt tctgcttttg catctttggg acacctcaga aaacttatta acaactgtga
1080
atatgagaaa tacagaagaa aataataagc cctctataca taaatgcccc gcacaattca
55 1140
ttgttaaaaa acaaccaaac ctacactac tgtatttcat tatctgtact gaaagcaaat
1200
gctttgtgac tattaatatg tgcacatcat tcattcactg tatagtaate attgactaaa
1260
60 gccatttgtc tgtgttttct tcttgtgggt gtatatatca ggtaaaaatat ttcccaaaga
1320
gccatgtgtc atgtaatact gaaccacttt gatattgaga cattaatttg taccctgtgt
1380
tattatctac tagtaataat gtaatactgt agaaatatgt ctctaattct tttcaaaatt
65 1440
gttgcacccc ccttagaatg tttctatttc cataaggatt taggtatgct attatccctt
1500

```



cttataccct aagatgaagc tgtttttgtg ctctttgttc atcattggcc ctcatcccaa  
 1560  
 gcactttacg ctgtctgtaa tgggatctat ttttgactg gaatatctga gaattgcaaa  
 1620  
 5 actagacaaa agtttcacaa cagattttct aagttaaatc attttcatta aaaggaaaaa  
 1680  
 agaaaaaaa tttttgtatg tcaataacct ttatatgaag tattaatatg catatttcta  
 1740  
 10 tgtttgtaata taatgagtca caaaataaag ctgtgacagt tctgttggc tacagaaatt  
 1800  
 tacttttgtg catttgtggc accacctact gttgaagggt tataaagcca ttagaaaagt  
 1860  
 agaggggaag tgatttggat caaaaggaaa aacttttagaa aagattcaaa tgttccctta  
 1920  
 15 atcataaaaag agaactgagg ggactacttg aaaataaaaag gttgttttgt attttcatgt  
 1980  
 tgggttaagat actgagtaac tgggtattaag tgttagaggt ttttagataa atattctgct  
 2040  
 20 taatgattat gaagctgcac tgagatttct gaaaatgctc tgtagctgag cttattttaa  
 2100  
 aaatgttcac ttggtatagg ggaagctaca aaggcagcct tcagtgtcct tttgtttatt  
 2160  
 caacaaaaaa tataaggaca caatgtagca gttatactgg gaaggtgctg ggggtggtg  
 2220  
 25 caatggtgag caggaaggcg aagtagatat ggaaacagaa atgatactaa ttcggtgat  
 2280  
 tccttccttt tttcctgtaa taagtgtgtg gcagacaaca tatgagcagt gctgataaat  
 2340  
 gtaaatgtat ttttcatagc tcattaagaa tcagtttcag aaagagatgt ctgcttattt  
 2400  
 30 tgctacttga agaaccctg tcaaacagtc cttttgagga agtacaagag gctgtctcta  
 2460  
 tttgtgacct caggaatggc tgtgacagtg tcgtgagcag tccttttcct gtggcacaga  
 2520  
 35 tctgaacttt gtgtgcagaa aaatcttggc ttcaagttag ccaagatgcc cctgagcat  
 2580  
 cagcatcaca acttcacct cctatcttga agttcatgtt atagtgaatt taatgaaatc  
 2640  
 atagaacact gtttcttcgt gaacaatgac gagggagagg aaaaaacttt attgaaaaat  
 2700  
 40 aaaaaggcag gtaatttaga tgaaaatatg ttacctatga ggttttgttt ttgctttttg  
 2760  
 tttttgtttt tgagaaacag aatctcgctc tgcgtccag gctggagtgc agcggcatga  
 2820  
 45 tcttggctca ctgcaacctc cgctcccg gttcaagcga ttctcctcag ctcccaagt  
 2880  
 agctggtact acaggcatgc gccaccacaa ccagctaatt tttgtatttt tagtagagat  
 2940  
 ggggtttcac tatacgttgg ccaggttgg ctcaaaactc tgacctaaagg tgatccttct  
 3000  
 50 gccttgggct cccaaagtgc tgggattaca ggcattgagc accttgctg gccctacca  
 3060  
 tgagccttga ctaaaacatt cttctatctg tagaaaagcc caaagaact tttccagatt  
 3120  
 55 caaaaaactt ggcactttgt aatggtaatg ttacattaa gtaaaaaaaa aaaaaaaaaa  
 3180  
 ccacttagc ttcagtttct aagtgtttac tgtgttgc tgcacttcat ttaatttca  
 3240  
 acacctgccc tatgaggtaa aaagtaccat ttacatatg agtaaattac agctcagtgg  
 3300  
 60 ataagaaact cgtccaaagg tacaggttca gtcaagtggc agagggttct tttgtttgaa  
 3360  
 gttaggatc agttaaaatt gacctgttaa aatcacatca gcatcaatat acattaattt  
 3420  
 65 aacaaatatt tattgaactt tactgtatgc cagatacttc tctaggtact aggggttaca  
 3480  
 atgtagaaga aaatagaatt c  
 3501

<210> 15  
<211> 151  
5 <212> ADN  
<213> Homo sapiens

<400> 15  
10 atgctgttgg acagcagggg caatgaccac ttttgggaac agcagttgga tggcttagat 60  
tggacagccc aagacatcgt ggcgtttttg gccaaagcacc cagaggatgt ccagtccagt  
120  
aatggttctg tgtacacctg gagagaagct t  
151

15  
<210> 16  
<211> 206  
<212> ADN  
<213> Homo sapiens

20  
<400> 16  
tgtgtcaacc tgaacaagct agaaccata gcaacagaag tctggctcat caacaagtcc 60  
atggagctgc tggatgagag gaagttctgg gctgggtattg tgttcaactgg aattactcca  
120  
25 ggcagcattg agctgcccc a tcatgtcaag tacaagatcc gaatggacat tgacaatgtg  
180  
gagaggacaa ataaatcaa ggatgg  
206

30  
<210> 17  
<211> 177  
<212> ADN  
<213> Homo sapiens

35  
<400> 17  
gtactgggac cctgggtctc gagctgaccc ctttgaggac atgcggtacg tctggggggg 60  
cttcgcctac ttgcaggatg tgggtggagca ggcaatcatc aggggtgctga cgggcaccga  
120  
40 gaagaaaact ggtgtctata tgcaacagat gccctatccc tgttacgttg atgacat  
177

45  
<210> 18  
<211> 223  
<212> ADN  
<213> Homo sapiens

<400> 18  
50 ctttctgctg gtgatgagcc ggtcaatgcc cctcttcatg acgtggcct ggatttactc 60  
agtggctgtg atcatcaagg gcatcgtgta tgagaaggag gcacggctga aagagaccat  
120  
gcggatcatg ggcttgagca acagcatcct ctggtttagc tggttcatta gtagcctcat  
180  
55 tctctctctt gtgagcgctg gcctgctagt ggtcatcctg aag  
223

60  
<210> 19  
<211> 222  
<212> ADN  
<213> Homo sapiens

<400> 19  
65 ttaggaaacc tgctgcccta cagtgatccc agcgtggtgt ttgtcttcct gtccgtgttt 60  
gctgtggtga caatcctgca gtgcttcctg attagcacac tcttctccag agccaacctg  
120

gcagcagcct gtgggggcat catctacttc acgctgtacc tgccctacgt cctgtgtgtg  
180  
gcatggcagg actacgtggg cttcacactc aagatcttcg ct  
222

5

<210> 20  
<211> 205  
<212> ADN  
10 <213> Homo sapiens

<400> 20  
agcctgctgt ctcctgtggc ttttgggttt ggctgtgagt actttgccct ttttgaggag 60  
cagggcattg gagtgcagtg ggacaacctg tttgagagtc ctgtggagga agatggcttc  
15 120  
aatctcacca cttcggcttc catgatgtg tttgacacct tcctctatgg ggtgatgacc  
180  
tggtagattg aggctgtctt tccag  
205

20

<210> 21  
<211> 15  
<212> ADN  
25 <213> Homo sapiens

<400> 21  
gccagtacgg aattc 15

30

<210> 22  
<211> 105  
<212> ADN  
<213> Homo sapiens

35

<400> 22  
gaattcccag gccctggtat tttncttgca ccaagtccta ctggtttggc gaggaagtg 60  
atgagaagag ccacctgggt tccaaccaga agagaatatc agaaa  
105

40

<210> 23  
<211> 132  
<212> ADN  
45 <213> Homo sapiens

<400> 23  
gtcaatcctg accgggttgt tccccccgac ctccgggcacc gcctacatcc tgggaaaaga 60  
cattcgctct gagatgagca ccatccggca gaacctgggg gtctgtcccc agcataacgt  
50 120  
gctgtttgac at  
132

55

<210> 24  
<211> 143  
<212> ADN  
<213> Homo sapiens

60

<400> 24  
gctgactgtc gaagaacaca tctggttcta tgcccgcttg aaagggtctt ctgagaagca 60  
cgtgaaggcg gagatggagc agatggccct ggatgttggt ttgccatcaa gcaagctgaa  
120  
aagcaaaaag agccagctgt cag  
65 143

<210> 25

<211> 138  
<212> ADN  
<213> Homo sapiens

5 <400> 25  
gtggaatgca gagaaagcta tctgtggcct tggcctttgt cgggggatct aaggttgta 60  
ttctggatga acccacagct ggtgtggacc ctactcccg caggggaata tgggagctgc  
120  
tgctgaaata ccgacaag  
10 138

<210> 26  
<211> 221  
15 <212> ADN  
<213> Homo sapiens

<400> 26  
gccgcaccat tattctctct acacaccaca tggatgaagc ggacgtcctg ggggacagga 60  
20 ttgccatcat ctcccatggg aagctgtgct gtgtgggctc ctccctgttt ctgaagaacc  
120  
agctgggaac aggctactac ctgaccttgg tcaagaaaga tgtggaatcc tccctcagtt  
180  
cctgcagaaa cagtagtagc actgtgtcat acctgaaaaa g  
25 221

<210> 27  
<211> 73  
30 <212> ADN  
<213> Homo sapiens

<400> 27  
gaggacagtg tttctcagag cagttctgat gctggcctgg gcagcgacca tgagagtgc 60  
35 acgctgacca tcg 73

<210> 28  
<211> 203  
40 <212> ADN  
<213> Homo sapiens

<400> 28  
atgtctctgc tatctccaac ctcatcagga agcatgtgtc tgaagcccg ctggtggaag 60  
45 acatagggca tgagctgacc tatgtgtgtc catatgaagc tgctaaggag ggagcctttg  
120  
tggaactctt tcatgagatt gatgaccggc tctcagacct gggcatttct agttatggca  
180  
tctcagagac gaccctggaa gaa  
50 203

<210> 29  
<211> 49  
55 <212> ADN  
<213> Homo sapiens

<400> 29  
atattcctca aggtggccga agagagtggg gtggatgctg agacctcag 49  
60

<210> 30  
<211> 114  
<212> ADN  
65 <213> Homo sapiens

<400> 30  
atggtacctt gccagcaaga cgaaacagge gggccttcgg ggacaagcag agctgtcttc 60

gcccggttcac tgaagatgat gctgctgac caaatgattc tgacatagac ccag  
114

5 <210> 31  
<211> 149  
<212> ADN  
<213> Homo sapiens

10 <400> 31  
aatccagaga gacagacttg ctgagtgga tggatggcaa agggctctac cagggtgaaa 60  
gctggaaact tacacagcaa cagtttggtg ccctttgtg gaagagactg ctaattgcc  
120  
gacggagtcg gaaaggattt ttgctcag  
15 149

<210> 32  
<211> 125  
20 <212> ADN  
<213> Homo sapiens

<400> 32  
attgtcttgc cagctgtgtt tgtctgcatt gcccttggtg tcagcctgat cgtgccaccc 60  
25 tttggcaagt acccagcct ggaacttcag ccctggatgt acaacgaaca gtacacatt  
120  
gtcag  
125

30  
<210> 33  
<211> 99  
<212> ADN  
<213> Homo sapiens

35 <400> 33  
caatgatgct cctgaggaca cggaaccct ggaactctta aacgccctca ccaaagaccc 60  
tggcttcggg acccgctgta tgggaaggaaa cccaatccc 99

40  
<210> 34  
<211> 189  
<212> ADN  
<213> Homo sapiens

45 <400> 34  
agacacgccc tgccaggcag gggaggaaga gtggacactg cccagttcc ccagaccatc 60  
atggacctct tccagaatgg gaactggaca atgcagaacc cttcacctgc atgccagtgt  
120  
50 agcagcgaca aaatcaagaa gatgctgcct gtgtgtcccc caggggcagg ggggctgcct  
180  
cctccacaa  
189

55  
<210> 35  
<211> 95  
<212> ADN  
<213> Homo sapiens

60 <400> 35  
agaaaacaaa aactgcaga tacccttcag gacctgacag gaagaaacat ttcggattat 60  
ctggtgaaqa cgtatgtgca gatcatagcc aaaag 95

65  
<210> 36  
<211> 33  
<212> ADN

<213> Homo sapiens  
<400> 36  
5 cttaaagaac aagatctggg tgaatgagtt tag 33  
<210> 37  
<211> 107  
<212> ADN  
10 <213> Homo sapiens  
<400> 37  
gtatggcggc ttttccttgg gtgtcagtaa tactcaagca cttcctccga gtcaagaagt 60  
15 taatgatgcc atcaaacaaa tgaagaaaca cctaaagctg gccaag 107  
<210> 38  
<211> 75  
<212> ADN  
20 <213> Homo sapiens  
<400> 38  
gacagttctg cagatcgatt tctcaacagc ttgggaagat ttatgacagg actggacacc 60  
25 aaaaataatg tcaag 75  
<210> 39  
<211> 170  
<212> ADN  
30 <213> Homo sapiens  
<400> 39  
gtgtgggttca ataacaaggg ctggcatgca atcagctctt tcctgaatgt catcaacaat 60  
35 gccattctcc gggccaacct gcaaaaggga gagaacccta gccattatgg aattactgct 120  
ttcaatcatc cctgaatct caccaagcag cagctctcag aggtggctct 170  
<210> 40  
<211> 178  
<212> ADN  
<213> Homo sapiens  
45 <400> 40  
gatgaccaca tcagtggatg tccttgtgtc catctgtgtc atctttgcaa tgtccttcgt 60  
cccagccagc ttgtcgtat tcctgatcca ggagcgggtc agcaaagcaa aacacctgca 120  
gttcatcagt ggagtgaagc ctgtcatcta ctggtctctt aattttgtct gggatatg 50  
178  
<210> 41  
<211> 116  
55 <212> ADN  
<213> Homo sapiens  
<400> 41  
tgcaattacg ttgtccctgc cacactggtc attatcatct tcattctgctt ccagcagaag 60  
60 tcctatgtgt cctccaccaa tctgcctgtg ctagcccttc tacttttget gtatgg 116  
<210> 42  
65 <211> 145  
<212> ADN  
<213> Homo sapiens

<400> 42  
 gtgggtcaatc acacctctca tgtaccacagc ctcttttggt ttcaagatcc ccagcacagc 60  
 ctatgtgggtg ctcaccacagc tgaacctctt cattggcatt aatggcagcg tggccacctt  
 120  
 5 tgtgctggag ctgttcaccg acaat  
 145

<210> 43  
 10 <211> 130  
 <212> ADN  
 <213> Homo sapiens

<400> 43  
 15 gggagaatcg ctttgtgtca ccattatctt gggacttggg gggacgaaac ctcttcgcca 60  
 tggccgtgga aggggtgggt ttcttctca ttactgttct gatccagtae agattcttca  
 120  
 tcaggcccag  
 130  
 20

<210> 44  
 <211> 121  
 <212> ADN  
 25 <213> Homo sapiens

<400> 44  
 acctgtaaat gcaaagctat ctctcttgaa tgatgaagat gaagatgtga ggcgggaaag 60  
 acagagaatt cttgatgggtg gaggccagaa tgacatctta gaaatcaagg agttgacgaa  
 120  
 g  
 121  
 30

35 <210> 45  
 <211> 63  
 <212> ADN  
 <213> Homo sapiens

40 <400> 45  
 atatatagaa ggaagcggaa gcctgctgtt gacaggattt gcgtgggcat tcctcctggt 60  
 gag 63

45 <210> 46  
 <211> 244  
 <212> ADN  
 <213> Homo sapiens

50 <400> 46  
 gtttgagat gggtatacaa tagttgtacg aatagcaggg tccaacccgg acctgaagcc 60  
 tgtccaggat ttctttggac ttgcatttcc tggaagtgtt ctaaaagaga aacaccggaa  
 120  
 catgctacaa taccagcttc catcttcatt atcttctctg gccaggatat tcagcatcct  
 55 180  
 ctcccagagc aaaaagcgac tccacataga agactactct gtttctcaga caacacttga  
 240  
 ccaa  
 244  
 60

<210> 47  
 <211> 1237  
 <212> ADN  
 65 <213> Homo sapiens

<400> 47  
 gaattcttca acagggaaaa cagctagctt gaaaacttgc tgaaaaacac aacttggtt 60

tatggcattt agtaccttca aataattggc ttgagagata ttggataccc cattaaatct  
 120  
 gacagtctca aatttttcat ctcttcaatc actagtcaag aaaaatataa aaacaacaaa  
 180  
 5 tacttccata tggagcattt ttcagagttt tctaaccag tcttattttt ctagttagta  
 240  
 aacatttgta aaaatactgt ttcactaata cttactgtta actgtcttga gagaaaagaa  
 300  
 aaatatgaga gaactattgt ttggggaagt tcaagtgate tttcaatata attactaact  
 10 360  
 tcttccactt tttccaaaat ttgaatatta acgctaaaag tgtaagactt cagatttcaa  
 420  
 attaatactt ctatattttt taaatttaca gaatattata taaccactg ctgaaaaaga  
 480  
 15 aaaaaatgat tgttttagaa gttaaagtca atattgattt taaatataag taatgaaggc  
 540  
 atatttccaa taactagtga tatggcatcg ttgcatttta cagtatcttc aaaaatacag  
 600  
 aatttataga ataatttctc ctcatltaat atttttcaaa atcaaagtta tggtttctc  
 20 660  
 attttactaa aatcgtattc taattcttca ttatagtaaa tctatgagca actccttact  
 720  
 tcggttctc tgatttcaag gccatattt aaaaaatcaa aaggcactgt gaactatttt  
 780  
 25 gaagaaaaca caacatttta atacagattg aaaggacctc ttctgaagct agaaacaatc  
 840  
 tatagttata catcttcatt aatactgtgt taccttttaa aatagtaatt ttttacattt  
 900  
 tcctgtgtaa acctaatgt ggtagaaatt tttaccaact ctatactcaa tcaagcaaaa  
 30 960  
 tttctgtata ttccctgtgg aatgtacctc tgtgagtttc agaaattctc aaaatacgtg  
 1020  
 ttcaaaaatt tctgcttttg catctttggg acacctcaga aaacttatta acaactgtga  
 1080  
 35 atatgagaaa tacagaagaa aataataagc cctctataca taaatgcccc gcacaattca  
 1140  
 ttgttaaaaa acaaccaaac ctccactac tgtatttcat tatctgtact gaaagcaaat  
 1200  
 gctttgtgac tattaaatgt tgcacatcat tcattca  
 40 1237

<210> 48  
 <211> 3002  
 45 <212> ADN  
 <213> Homo sapiens

<400> 48  
 acatggcaat ggcattcatt aggaatctag ctgggaaaat ccagtgtgta tgcttgaaa 60  
 50 tgagggatct ggggctggag agaaaggcat gggcatgcct tggagggact tgtgtgtcaa  
 120  
 gctgaggacc tttactttaa gctctagggg accaggcaag gggagatgta gatacgttac  
 180  
 tctgatggg tggatgaatt gaagaaggat gaggcaagaa tgaaggcaga gaccagggag  
 55 240  
 gaggctctcc aagtggccaa ggcataaagc aagaaatgag gcttggtgac tgcttagtgg  
 300  
 cagagcagtg aaagagaggg aggcatacaa gtgagtctcg atttctagct ggggtgggtg  
 360  
 60 tagcgatgtc cagtaggcca gtggctactg aggtctgcag tggaggaggg tgggtgggct  
 420  
 ggagacagat gatgaggag tcatcagcct gtgggtggaa gaaaaggga cctcttccaa  
 480  
 ctgttttctt tgcttcttcc ctctcttctc cttttttttt ttttttgac agagtcttgc  
 540  
 65 tctgtcacc aggtgaaat gcagtggcat gatcttggt caccacagcc tccgcctcct  
 600



gggttcaagc aattctcctg tctcagcctc cagagtagct gggattacag gcacatatca  
 660  
 ctgtgcccg ctaatttttg tattttcagt ggagatggga ttccaccatg ttggtcgggc  
 720  
 5 tggaatgaac tcttgacctc aagtgatcca cctgcctcag cctcccaaag tgttgggatt  
 780  
 acaggcattg agccaccgag cccggccttt cttccctctc ttaaagagtg tttatttaac  
 840  
 10 tccacaaaca tgagcttgtc accccctgta gcctggcatc tcttacacga ggtgatggct  
 900  
 gaggtctctg cttctgctgg ggtagctctg atctttctgc tttctctggc actgtctacc  
 960  
 catgttgctt caccaccag gtcccagggc acctctctcg ggcaagtctt ggaacctctt  
 1020  
 15 gacactgatt tgctctcttt tctgagctgc ttttagccac ccatcctcgg gacctgtttt  
 1080  
 ctctctgctt ccacctctgc gggcagctct aggtctcctg cccctcacga gcacccaga  
 1140  
 gaggccactg gctcagtgat ctgagtgagg gcatctttct agtcttgcta ttttttttgg  
 1200  
 20 ccattgttgt cagaaacct actgggcagg gccgacttca ccctaaaggc tgcgtctctt  
 1260  
 cactctgctt ttgtttgttc caataaagt ggcttcagaa ttgctaacc tagcctctgt  
 1320  
 25 gaactttgtg ggtacaattt tgtgtctgtt atgttaacaa aaatacatat atacctctt  
 1380  
 ggtgatggta taaattgcta ttctctattg gaaagcaatt tggaatgaaa atttaagaa  
 1440  
 30 ccattttaaa atatgctatc ctgcgtacct ccattccacc cacccccagg gatgtagcct  
 1500  
 actgaaataa ttttaagaa gtcaccatat gagagaaaat gttattgcta tattgttatt  
 1560  
 gtgagaaatt ggaaatagac taaatgttca gcactatagg aataattaat gaaattacat  
 1620  
 35 atactctata caatcattat gctgccattg aaataataaa taaaaggcgg caagggggga  
 1680  
 aaagcttata atgttagtga aactaagact gattttttta taaagcagca gttttcagac  
 1740  
 40 ccttgagagac tccaattcgg tagaaccaga gcttcactct ctctgtcgaa gctgtgacag  
 1800  
 gagttgcaaa tgcctctcct ttttgctgag ttgagcgtg ctgtttttcc ggcagcacat  
 1860  
 ctgtgcaggc ctctgcctcg gcccctctgg atctgctgat tgagcagcgg attgatctgt  
 1920  
 45 ccttctcttt cgtgttgacc catgtgagga accaactggc aagggaacaa gaaatggaaa  
 1980  
 taggcctcct ttgcatcatg acctgtacat cctgcaattg gaaaagattg tacttttagt  
 2040  
 ggtttaacca gcagcattat ttttctaaac taagcagtaa gaaggaaatta ggttttatgt  
 2100  
 50 gggatcaaca gactgggtct caaaagagga aggtgataga acacagtggg gagggggagg  
 2160  
 tgcactagaa acagagggcc tatgctttca ttctggcttt gctacttaac agctgtgtga  
 2220  
 55 cccaatctta gagacttaac ctctctgaac ttccattttc tcatgtataa aatgggaaat  
 2280  
 attaaaggat actcactggg ctggtggctt gtgcctgtaa tcccagcact tggggagggt  
 2340  
 gaggtgggag gatcacttga gcccagggtg tcaagaccag cccaggcaac atggcaagac  
 2400  
 60 tctgtctcta tgaaaaaatt aaaaattagc caggtgtggt ggtgtgcacc tgtagtctta  
 2460  
 gctacttggg aggtgagat gggaggatca ctggggcttg ggaggtcaag gctgcggtga  
 2520  
 65 gctgtgatcc catcactgca ctccagcccg ggcggcagag cgagacactg aatccaaacg  
 2580  
 acaacaacaa caaaaggcaa aaaaataaaa gtgcctctt tatggagttg tgaagggtga  
 2640

agcatatata ctattcaaca tagtaactat ataaaggaag tattgttggt gttactgtag  
 2700  
 ttaataccat taagtgaat gtttcgtata gtggaaagca catggactct gaattcagac  
 2760  
 5 tgggtctgact ttgagtctca gctccacatc tagtaatact atgaccaagc cctgggttaa  
 2820  
 atcatgtttt tttttcttca gccttagtct tctcacatat aaaatagga cactgtcatt  
 2880  
 tacctcagtt ttctgtgagg ataaaacaac gacagtgtat atgcaagtat tttgtaaatt  
 10 2940  
 ttgtagtgt cctcaagatt tagttgggtt ttactacttg tactttctca ctggaatggc  
 3000  
 ag  
 3002  
 15  
 <210> 49  
 <211> 397  
 <212> ADN  
 20 <213> Homo sapiens  
 <400> 49  
 aattcgggtc caattaaatt tttgaaattt tatattaaaa attatattag tagggatggg 60  
 taagagggtg tttgggtctg ttgggttggt agttgctatg actcagaatt gctaagaaaa  
 25 120  
 cagaaaagta agataagatc attgttttaa cctcttttcc tccacaaaat caataaataa  
 180  
 catatcccta aattactctt agaatttctc ttaaattgca gtgaaaaacc aaaatccttc  
 240  
 30 attcttggtt gaagggttga aaactacgtt agagaggatt agagagagag gatgagcaat  
 300  
 cgtgtagtca gcccttgccct cctagtgtag gatttgtctc agccactgct tgttgtcctg  
 360  
 gctgccaacg ttctcatgaa ggctgttctt ctatcag  
 35 397  
 <210> 50  
 <211> 520  
 40 <212> ADN  
 <213> Homo sapiens  
 <400> 50  
 gtaagtggaa tcccatcaca ccagcctggt cttggggagg tccagagcac ctattatatt 60  
 45 aggacaagag gtactttatt ttaactaaaa atttggtaga aatttcaaca acaacaaaaa  
 120  
 aactcaactt ggtgtcatga ttttggtgaa attggtacat gacttgctgg aagggttttc  
 180  
 ataggtcata aaataacagt atcttttgat tttagcatttc tactcaaggg aattaattcc  
 50 240  
 aggaattttg gtggcaggca cctgtaatcc cagctactcg ggaggctgag gcaggagaat  
 300  
 tgcttgaacc caggaggcag aggttgagc gagctaagat cgcacattg cactcccgcc  
 360  
 55 tgggcaataa gagtgaact ccactcaca aaaaaaaaaa gatacaaaaa tagaaaaagg  
 420  
 ggcttggtta gggtagtagg gttttgggca attttttttt tttttttttt ttattgtatg  
 480  
 gttctaaagg aatgggtgat tacctgtggt ttggttttag  
 60 520  
 <210> 51  
 <211> 1786  
 65 <212> ADN  
 <213> Homo sapiens  
 <400> 51

gtaagttacc tgcaagccac tgtttttaac cagtttatac tgtgccagat gggggtgtat 60  
 atatgtgtgt gcatgtgcat gcatgtgtga atgatctgga aataagatgc cagatgtaag  
 120  
 5 ttgtcaacag ttgcagccac atgacagaca tagatatatg tgcacacact agtaaaccctc  
 180  
 tttccttctc atccatgggt gccactttta tctttttatt tttatttttt tttttgagat  
 240  
 ggagtctcgc tctgacgccc aggctggagt gcagtggctc gatctcggct cactgcaacc  
 300  
 10 tttgcctccc gggttcaagc tattctcctg cctcagcctc cacagtagct gggactacag  
 360  
 gctcatgctg ccacgcccgg ctgacttttt gtatttttagt agagacgagg tttcaccatg  
 420  
 ttaccaggc tagacttcaa ctctgagct caggcaatcc accctccttg gcctcccaaa  
 480  
 15 gtgctgggat tacaggtgtg agccactgca ccagccac cactttaatt ttttacctc  
 540  
 tacccttttg gtcaaaattt gctcaatctg caagcttaaa atgtgtcatg acaaacacat  
 600  
 20 gcaagcacat actcacacat agatgcagaa acagcgtcta aacttataaa agcacagttt  
 660  
 atgtaaatgt gtgcaattct tctccctagg tggtaaacca catttcaaaa caacccaat  
 720  
 25 aaaaactgaac aaagcttctt cctcttagac tttttagaaa atctttcagt gctgagtcac  
 780  
 taagctgcca agttctcatt gtgggaacta tgcctttgga tgtaatgatt tcttctaaga  
 840  
 caatgggchg aggtgtagtt attgcagaca tctgaaatat gtaatgtttc ttccagattc  
 900  
 30 tggaaattct cttattctct gtggttggtg gtggttggtg gatgtgtgtg tgtgtgtgtg  
 960  
 tgtgtgtgtg tgtgtgtgtg tagggatcag gatgchgag gagctgggtt ctgcttgtat  
 1020  
 tggttctctg ttttgcattg aatagtgtgt ttccttgtat ggctatctat agcttttcaa  
 1080  
 35 ggtcaccaga aattatcctg tttttcacct tctaaacaat tagctggaat ttttcaaagg  
 1140  
 aagactttta caaagacccc taagctaagg tttactctag aaaggatgtc ttaagacagg  
 1200  
 40 gcacaggagt tcagaggcat taagagctgg tgcctgttgt catgtagtga gtatgtgcct  
 1260  
 acatggtaaa gctttgacgt gaacctcaag ttcagggtcc aaaatctgtg tgccttttta  
 1320  
 ctttgacat ctgcattttc tattctagct tggaaatctga aacattgaca agagctgcct  
 1380  
 45 gaaatgtatg tctgtggtgt gattagagtt acgataagca agtcaatagt gagatgacct  
 1440  
 tggagatggt gaacttttgt gagagaatga gttgtttttt tgttttggtt tttagtactt  
 1500  
 50 taacataatc taccttttagt ttaagtatcg ctcacagtta cctagttact gaagcaagcc  
 1560  
 cccaaagaaa tttggtttgg caacactttg ttagcctcgt ttttctctct acattgcatt  
 1620  
 gctcgtgaag cattggatca tacgtacatt tcagagtcta gagggcctgt ccttctgtgg  
 1680  
 55 cccagatgtg gtgtccctc tagcatgcag gctcagaggc cttggcccat caccctggct  
 1740  
 caggtgtgtc tttctttctc cccttgcct tcttggggc ctccag  
 1786  
 60  
 <210> 52  
 <211> 1745  
 <212> ADN  
 65 <213> Homo sapiens  
 <400> 52  
 gtaaggcagc ctcaactcgt cttccctgcc aggaaactcc gaaatagctc aacacgggct 60

aagggaggag aagaagaaaa aaaatccaag cctctggtag agaaggggtc atacctgtca  
 120  
 ttctctgcaa ttatcatccat ttatagttgg ggaaagttag gccagagag gggcagtgac  
 180  
 5 ttgcccagg tcaaccacgc cgggtagcag ctaagtagga tgagagtga gggttcatgc  
 240  
 ttccagata accacatgct caactgtgcc atgctgtctc attggtagt gttcatggca  
 300  
 10 gcatctgaaa gctatttatt ttcttagata tattgggtgg cgattcttcc taagtttcta  
 360  
 agaacaataa tcagaaggat atatattgtt gcagggttaga ctgtctggaa gcagaggctg  
 420  
 aaatagagtt tgatgtatgg gtatttatga gggctcaata cctatgaaga gatatggaag  
 480  
 15 atgcaggatt gggcagaggg aggagttaga ctgtgatata gggccaaccc cgtggggcac  
 540  
 tctanagaat atgcagcttg ttggagttgt tnttcatcga gctgaaacat ccagcccttt  
 600  
 gtgctcccc aaggcctccc tctgacacc acctacctca gccctctcaa tcaatcactg  
 660  
 20 gatgtgggt gccctgggaa ggtcgtgccc cagggcctac atggctctct gctgctgtga  
 720  
 caaaccaga gttgctgatg cctgaggccg tctactgaca gctgggcaac aaggcttccc  
 780  
 25 tgaatgggga ctctgggcag tgcagttttg tgtctgaacc atacattaat atatttatat  
 840  
 ccgaattttc ttctctgca agcatttcat ataaagacac atcaggtaaa aataaatgtt  
 900  
 30 ttgaagcaa aaggagtaca aagagataag aactaactaa ttaatacta gttaccatct  
 960  
 gttacaaata gttctactg attgccaagg actgttttaa cacatcacat gggcttcttc  
 1020  
 ttctatcctc actaaccctt ttaacagaca aggaaatgag gctcaggaag gtcaaggact  
 1080  
 35 ttattgaggt tccacagtag gatacagttc ttgctaaaag caaccctcc ctcatgctct  
 1140  
 gttatctaac tgcaagggga aggtcagtg cagaggtagt ggtcccatgg ttggtgcata  
 1200  
 agagctgctc tgagacaact gcatgctggg ggtcctgca gacatgtacc catcagccgg  
 1260  
 40 agataggctc aaaatatcca caagagtttg gatgattgtg ggaatgcaga atccatggtg  
 1320  
 atcaagaggg aaagtcaagt tgctggcca ttttcttgg ctttttagaca gaaaagttac  
 1380  
 45 gtgggatatt atctcccaca gctcttctgt ggtgccacca gtcatagtcc ttatataagg  
 1440  
 agaaaccagt tgaaattacc tattgaagaa acaaagagca aactcgcca ctgaaatgcg  
 1500  
 tagaaagccc tggactctgt tgtattcata actctgccat ttttttctg cgtagttttg  
 1560  
 50 ggtaagtcac ttatcttctt taggatggta atgatcagtt gcctcatcag aaagatgaac  
 1620  
 agcattacgc ctctgcattg tctctaacat gagtaggaat aaaccctgtc ttttttctgt  
 1680  
 55 agatcataca agtgagtgtc tgggattgtt gaggcagcac atttgatgtg tctcttctt  
 1740  
 cccag  
 1745  
 60  
 <210> 53  
 <211> 1060  
 <212> ADN  
 <213> Homo sapiens  
 65  
 <400> 53  
 gtgagtacct ctggccttct ttcagtggct gtaggcattt gaccttctt tggagtcctt 60

gaataaaagc agcaagttga gaacagaaga tgattgtctt ttccaatggg acatgaacct  
 120  
 tagctctaga ttctaagctc ttttaagggtta agggcaagca ttgtgtttta ttaaattggt  
 180  
 5 tacctttagt cttctcagtg aatcctgggt gaattgaatt gaatggaatt tttccgagag  
 240  
 ccagactgca tcttgaactg ggctggggat aaatggcatt gaggaatggc ttcaggcaac  
 300  
 10 agatgccatc tctgcccttt atctcccagc tctgttggct atgttaagct catgacaaag  
 360  
 ccaaggccac aaatagaact gaaaactctt gatgtcagag atgacctctc ttgtcttctt  
 420  
 tgtgtccagt atgggtgttt gcttgagtaa tgttttctga actaagcaca actgaggagc  
 480  
 15 aggtgcctca tcccacaaat tcttgacttg gacacttctt tccctcgtac agagcagggg  
 540  
 gatattcttg agagtgtgtg agccccatac agtgcaagtt gtcagatgtc cccagggtcac  
 600  
 ttatcaggaa agctaagagt gactcatagg atgctcctgt tgcctcagtc tgggcttcat  
 660  
 20 aggcacagc agccccaaac aggcacctct gatcctgagc catccttggc tgagcagggg  
 720  
 gcctcagaag actgtgggta tgcgcatgtg tgtgggggaa caggattgct gagccttggg  
 780  
 25 gcactcttgg aaacataaag ttttaaaagt tttatgcttc actgtatatg catttctgaa  
 840  
 atgtttgtat ataatgagtg gttacaaatg gaatcatttt atatgttact tggtagccca  
 900  
 ccactccctt aaagggactc tataggtaaa tactacttct gcaccttatg attgatccat  
 960  
 30 tttgcaaatt caaatttctc caggtataat ttacactaga agagatagaa aaatgagact  
 1020  
 gaccaggaaa tggataggtg actttgcctg tttctcacag  
 1060  
 35  
 <210> 54  
 <211> 1104  
 <212> ADN  
 40 <213> Homo sapiens  
 <400> 54  
 gtacactgct ttgggcatct gtttggaaaa tatgacttct agctgatgtc ctttctttgt 60  
 gctagaatct ctgcagtgca tgggcttccc tgggaagtgg tttgggctat agatctatag  
 120  
 45 taaacagata gtccaaggac aggcagctga tgcgtgaaagt acaattgtca ctacttgtac  
 180  
 agcacttggt tcttgaaaac tgtgtgccag gcagcatgca aaatgtttta tacacattgc  
 240  
 50 ttcatttaat tctcacaagg ctactctgaa gtagttaacta taataaccag caattttcaa  
 300  
 atgagagaac tgtgactcaa agacgttaag taaccagctt tggtcacaca actgttaaat  
 360  
 gttgggtacgt ggaggtgaat ccacttcggt tacactgggt caataagccc aggcgatcc  
 420  
 55 tcccaatgct cacccaattc tgtatttctg tgtcctcaga gggggtacaa ctaggagag  
 480  
 tttgtgttcc tgagtacagg ttgttaataa ttaaataatac tagctctaag gcctgcctgt  
 540  
 60 gatttaatta gcattcaata aaaattcatg ttgaattttt ctttagtact tctttcttaa  
 600  
 tataatacat cttcttgacc aagtccaaga ggaacctgcg ttggacagtt ttcatatgag  
 660  
 atcaaattct gagagagcaa gatttaaccc tttttggttc accttctgat cctcccctaa  
 720  
 65 ggaggtatag atgaaatatt tattactcct gcctgaactt ctttcattga atatgcaatt  
 780

ttgcagcatg cagattctgg atttaaattc tgagtcctaa cttactggct gagggacctt  
840  
ggataggctc cttatccctc agtttctcctc tctctaaaaat ggggatggca cctgccccgt  
900  
5 ggggtgttgg aaggacttac agagggtgag aatgtacgtt gtacatagca ggtttcagca  
960  
aatgttagct cctcttttcc ccacatccat tcaaatctgt tccttctcca aaggatgtgt  
1020  
10 caaggaggaa atggacctgg ctgggaaacc ctcagaatac tgggatgatg ctgagcttgg  
1080  
ctcatacctg tgctttgctt tcag  
1104

15 <210> 55  
<211> 1180  
<212> ADN  
<213> Homo sapiens

20 <400> 55  
gtaagtgtg ttgacctcct gctctttctt taacctagtg ctgctgcctc tgctaactgt 60  
tgggggcaag cgatgtctcc tgcctttcta aaagactgtg aaaccactcc aggggcagag  
120  
aaatcacatg cagtgtccct ttccaaatcc tcccatgccca tttatgtcca atgctgttga  
25 180  
cctattggga gttcacgggc tcgatccctg agggacattt tctttgttgt cttggcttct  
240  
agaagagtat cttttacttg cccctcccca aacacacatt tcatggctc ctaacaagct  
300  
30 agaagaaaga ggtaaagaca agcgtgattg tggaaaccata gcctcgctgc ctgcctgtga  
360  
catggtgacc tgtgtatcag cctgtgtggg ctgagaccaa gtggctacca cagagctcag  
420  
cctatgcttc ataattgaat cattaccagc atccctaate ctctcttggc tcttaactgc  
35 480  
agacagagat gtccacagct catcaaaggc tctgccttct gggttctttg tgcttagagt  
540  
ggcttctctaa atatttaata ggtccctttt ctgccagtct cttctgtgcc catccctga  
600  
40 ttgcccttgg taaaagtatg atgcccctta gtgtagcagc cttgcctgct gttcctaate  
660  
atcttctcct acctcctctt tacacctagc tcctgtttca gtcacctaga aatgtcaca  
720  
gtcgctggaa tatgtcatgt tcttccacac ctccatgcct ttgtaggtac tgtttgtctt  
45 780  
cacaggagaa ctttctctct aacttgccca tcttctcaac tcctcctttc tctccaagat  
840  
ctagttccgg atccctccc ctgagcatcc ctcttgggt ctcaggtagt cagtcactct  
900  
50 ctgccctgaa cttccatggc acgtgaaaga aaatctttt attttaaaac aattacagac  
960  
tcacaagaag taatacaaat tacatgaggg ggttccctta aacctttcat ccagtttccc  
1020  
caatggtagc agcatgtgta actgtagaat agtatcaaaa ccatgaaatt gacataggta  
55 1080  
caattcaca accttcttca gatttacta gctttatgtg cgctcatttg tgtgtgtgtg  
1140  
tgcgtattta gttctatgca attttatcat gtgtgaattc  
1180

60

<210> 56  
<211> 903  
<212> ADN  
65 <213> Homo sapiens

<400> 56  
gatccctggg ccaagggaag gagcacatga ggagttgccg aatgtgaaca tgttatctaa 60

tcatgagtgt ctttccacgt gctagtttgc tagatgttat ttcttcagcc taaaacaagc  
 120  
 tggggcctca gatgaccttt cccatgtagt tcacagaatt ctgcagtggc cttggaacct  
 180  
 5 gcagccacga aaagatagat tacatatgtt ggagggagtt ggtaattccc aggaactctg  
 240  
 tctctaagca gatgtgagaa gcacctgtga gacgcaatca agctgggcag ctggcttgat  
 300  
 10 tgccttccct gcgacctcaa ggaccttaca gtgggtagta tcaggagggg tcagggggctg  
 360  
 taaagcacca gcgttagcct cagtggcttc cagcacgatt cctcaacat tctaaccatt  
 420  
 ccaaagggtat tatctttggg ggggtgacatt cttttcctgt tttcttttta atcttttttt  
 480  
 15 aaaacataga attaatatat tatgagcttt tcagaagatt tttaaaggc agtcagaaat  
 540  
 cctactacct aacacaaaaa ttgtttttat ctttgaataa tatgttcttg tttgtccatt  
 600  
 20 ttccatgcat gcgatgttag gcatacaaaa tacatttttt aaagaatact ttcattgcaa  
 660  
 attggaaact tcgtttaaaa aatgctcata ctaaaattgg catttctaac ccataggccc  
 720  
 acctgtagtt atttaccgaa gcaaaaggac agctttgctt tgtgtgggtc tggtaggggt  
 780  
 25 cattagaaag gaatgggggc ggtggggagg ttggtgttct gttctctctg cagactgaat  
 840  
 ggagcatcta gagttaagg taggtcaacc ctgacttctg tacttctaaa tttttgtcct  
 900  
 cag  
 30 903

<210> 57  
 <211> 486  
 35 <212> ADN  
 <213> Homo sapiens

<400> 57  
 40 gtgagtacca gcagcacgtt aagaataggc cttttctgga tgtgtgtgtg tcatgccatc 60  
 atgggaggag tgggacttaa gca:tttact ttgctgngtt tttgtttttt ctttttttct  
 120  
 tttttttttt tttgagatgg agtctcgctc tgtagccagg ctggactgta gtggcgcat  
 180  
 45 ctcggtcac tgcaaccttg gcctcccagg ttcaagcat tctctgcct cagcctcccg  
 240  
 agtagctggg actctaggca cacaccacca tgcccagcta atttttgtgt ttttagtaga  
 300  
 gacgggggtt caccatgttg gccaggatgg tctcaatgtc ttgacctgt gatccgccca  
 360  
 50 cctcgggtct ccaaagtgtt gggaacacag gcatgagcca ctgtgtctgg ccacatttta  
 420  
 ctttctttga atatggcagg ctacacctcg tgaacacctt gagacctagt tgttctttga  
 480  
 ttttag  
 55 486

<210> 58  
 <211> 283  
 60 <212> ADN  
 <213> Homo sapiens

<400> 58  
 65 gaaattgaaa gttgtaactg cctgggtgcat ggtggccagg cctgctggaa acaggttggg 60  
 agcgatctgt cacctttcac ttgatttcc tgagcagctc atgtggttgc tactgttgt  
 120  
 tctaccttga atcttgaaga ttatttttca gaaattgata aagttatttt aaaaagcacg  
 180

gggagagaaa aatatgccca ttctcatctg ttctgggcca ggggacactg tattctggg  
 240  
 tatccagtag ggcccagagc ttgacctgcc tccctgtccc cag  
 283  
 5  
 <210> 59  
 <211> 203  
 <212> ADN  
 10 <213> Homo sapiens  
 <400> 59  
 gtgcggccca gagctacctt ccctatccct ctcccctcct cctccggcta cacacatgcg 60  
 gaggaaaatc agcactgccc cagggtccca ggctgggtgc ggttggtaac agaaacttgc  
 15 120  
 ccttggtgtg gccctaggt cctctgcctt cactcactgt ctggggctgg tcctggagtt  
 180  
 tgtcttgctc tgtttttttg tag  
 203  
 20  
 <210> 60  
 <211> 702  
 <212> ADN  
 25 <213> Homo sapiens  
 <400> 60  
 gtgcctgatg tgtatttatt ctgagtaaat ggactgagag agagcggggg gcttttgaga 60  
 agtgtggctg tatctcatgg ctaggcttct gtgaagccat gggatactct tctgttatca  
 30 120  
 cagaagagat aaagggcatt gagactgaga ttcttgagag gagatgctgt gtctttattc  
 180  
 atctttttgt cccaacatg gtgcactaaa tttatggtta gttgaaaggg tggatgctta  
 240  
 35 aatgaatgga agcggagagg ggcaggaaga cgattgggct ctctggttag agatctgatg  
 300  
 tggtagagta tgaggagcac aggcaggctt ggagccaact ctggctggcc ctgagacatt  
 360  
 gggaaagtca caacttgcct caccctcttt gccgataata atagtgtgtc ttacctcata  
 40 420  
 gaggattaaa ttaaatgaga atgcacacaa accacctagc acaatgcctg gcatatagca  
 480  
 agttcccaaa taaatgcta ctgttcttac ctctgtgagg atgtggtacc tatatataca  
 540  
 45 aagctttgcc attctagggg tcatagccat acagggtgaa aggtggcttc caggctctct  
 600  
 ccagtgtta cccctgctaa tatctctcta gtccctgtca ctgtgacaaa tcagaactga  
 660  
 gaggcctcac ctgtcccaca tccttgtgtt tgtgcctggc ag  
 50 702  
 <210> 61  
 <211> 1258  
 55 <212> ADN  
 <213> Homo sapiens  
 <400> 61  
 gtgagctgca gtcttggtgc tgggtgtgtg ttgggtctgg gcagccagga ctgctggt 60  
 60 gtgaatgatt tctccatctc caccctttt gccatgttga aaccaccatc tccctgctct  
 120  
 gttgcccctt tgaatcata tcatacttaa ggcattgaaa gctaaggggc cctctgctcc  
 180  
 cattgtgcta gttctgttga atcccgtttt ccttttcta tgaggcacag agagtgtg  
 240  
 65 agaaggtcct tagaggacat tattatgtca aagaaaagag acttgtcaag aggtaagagc  
 300



ctgggtaca aatgacctgg tggctctgct cattactttt caatctcatt gaccttaact  
 360  
 tttaaactat aaaacagcca atatttatta ggcactgatt tcatgccaga gacactctgg  
 420  
 5 gcatgaaaga aagtaatgat aatagttaat tttatatagc gttgtttacca tttaacaact  
 480  
 tttttttttt tttaacctct atcatctcaa ttaaagtga gagagaccct gggaagaagg  
 540  
 taactatatt tattatccca gatgagggaa gtgaggcttg tagggaattg gttagctgatt  
 10 600  
 caaggtcacc cagcaggtaa ataacagtgg tgggaccaga cccaattacc aggtatgttt  
 660  
 tcctctgtac cgcagtacat gcctgagatt tatttgtgtg ttgaagccag tggtagctaa  
 720  
 15 tgtatttaca tcccaacctg aaactcctat ccacttattt acccttttaac gagcctctta  
 780  
 actcaagtgc agtctgagga ccagcagcat caggatcact tgggaacttg ttagaaattc  
 840  
 agcaacctgg gccagctca gacctaccga atcagaatct gtgcatttta acaaggttct  
 20 900  
 tgagtgttg aacacacatt aaagcatgag aagcattgaa ctagacatgt agccaggtaa  
 960  
 aggccttgcc tgagatggtt ggcaaaggcc tcattgcage attcattggc aggccacagt  
 1020  
 25 tcttttgga gctctgctc ctgaccttc accctcagga agcgaggctg ttcacacggc  
 1080  
 acacacatgc cagacagggt cctctgaagc cacggctgcc agtgcatgtg tcccaggga  
 1140  
 agctttttcc ttagttctc acacaacaga gcttcttga agccctcccc ggcaaagggt  
 30 1200  
 ctggtggctc tgccttgctc cgtccctgac ccgttctcac ctcttcttt gccatcag  
 1258  
  
 35 <210> 62  
 <211> 986  
 <212> ADN  
 <213> Homo sapiens  
  
 40 <400> 62  
 gtaaggactc tggggtttct tttcaggtg gtgcctgagc ttccccagc tgggcagagt 60  
 ggaggcagag gaggagaggt gcagaggctg gtggcgctga ctcaaggttt gctgctgggc  
 120  
 tggggctggg tggctgcggg tgtgggagca gcttgggtggc ggggtggcct aatgcttgct  
 45 180  
 ggggtgcctg gggctcgggt tgggagctag cagggcagtg tcccagagag ctgagatgat  
 240  
 tggggtttgg ggaatccctt aggggagtg acactgaata ccagggatga ggagctgagg  
 300  
 50 gccaaagccag gagggtggga tttgagctta gtacataaga agagtgagag cccaggagat  
 360  
 gaggaacagc cttccagatt tttcttgggt agcgtgtgta ggaggccagt gtcaccagta  
 420  
 gcatatgtgg aacagaagtc ttgaccttg ctatctctgc ctagtccctaa tggctggctt  
 55 480  
 ttcccaggaa ggcttctgct tncatggacn gntagattaa ccctttattt aggtaaatga  
 540  
 gggaaacctac tttataagca taggaaaggg tgaagaatct tttaagattc ctttactcaa  
 600  
 60 gttttctttt gaagaatccc agagcttagg caatagacac cagactttga gcctcagtta  
 660  
 tccattcacc catccacca cccaaccacc catccttcca tcctcccatc ctcccattca  
 720  
 cccatccacc catccagctg tccaccatt ctacactgag tacctataat gtgcctggct  
 780  
 65 ttggtgatac aaaggtgaat aagacatagt cctttccttt gcccacaacc ctccagaccag  
 840

agatgaacat gtggaatgac ctaaacacct ggaacagggtg tgggtgtatga gcggcaggcc  
 900  
 tctgatgaga ggggtgggga tggccagccc tcaactccgaa gccctctga gttgattgag  
 960  
 5 ccatctttgc attctggtcc ctgcag  
 986

<210> 63  
 10 <211> 1667  
 <212> ADN  
 <213> Homo sapiens

<400> 63  
 15 gtaagttaag tggctgactg tcggaatata tagcaaggcc aaatgtccta aggccagacc 60  
 agtagcctgc attgggagca ggattatcat ggagttagtc attgagtttt taggtcatcg  
 120  
 acatctgatt aatgttggcc ccagtgcagc atttaagatg gtagtgggag atagcaggaa  
 180  
 20 agaagtgttt tcctctgtac cacagtacat gcctgagatt tgtgtgttga aaccagtggg  
 240  
 acctaacaca ttacatccc aaccttaaac tcctatgcac ttatttacc tttaatgagc  
 300  
 ctctttactt aagtacagtg tgaggaacag cggcatcagg atcacttggg aacttgttag  
 360  
 25 aaatacagca acttgggccc agctcagacc tactgaatca gaatcaggag caattctctg  
 420  
 gtgtgactgt gtcacagcca ggtatcaact ggattctcat acataggaaa tgacaaacgt  
 480  
 30 ttatggatgg atagtctact tgtgccagggt gctgagattt gttttttgtt ttttgatttt  
 540  
 ttttaataca ctgtgacctc atttaattct caaaaaaaga tgaaaaaatg aacactcagg  
 600  
 aatgctgaca tgagattcag aatcaggggt ttggggcttc aaagtccatc ctctctttat  
 660  
 35 ccatgtaatg cctccctta gagatacaac atcacagacc ttgaaggctg aaggggatat  
 720  
 aaaagctgtc tggccaagtg gtctccaagc ttgacagtgc agcagaatca cctggggata  
 780  
 40 ttattaaaaa taaacatact aaggtttggc ttcagggcct gtgaatcaga atttctggag  
 840  
 gtgaggcctt gaagtctgta tttctattgc atactttgga cacagtgggc tatagactag  
 900  
 agttttgaaa tgattgcgct cattcagatt ctcttctgat gtttgaattg ctgccatcat  
 960  
 45 atttctagtg ctctatttcc tcctgctcat tctgtcttgg ataacttatc atagtactag  
 1020  
 cctactcaaa gatthagagc cacagtctg aaagaagcca cttgactcat tccctgtagg  
 1080  
 50 ttcagaataa atttcttctg cgcagtgtct gtcatagctt tttttaaat tttttttatt  
 1140  
 tttgatgaga ctggagtttt gctcttattg cccaagctgg agtgagtggtg tgcgattttg  
 1200  
 gctcactgca acctccacct cccaggttca agcgattctc ctgcctcagc ctcccaagta  
 1260  
 55 gctgagatta caagcatgtg ctaccacgcc cagctaattt tgtattttta gtagagatgg  
 1320  
 gttttatcca tgttggctag gctggctctg agctccagac ctgaggtgat ctgcccgcct  
 1380  
 60 cggcctccca aagtgtggg attataggcc tgagccacag cgctcagcca taactttaat  
 1440  
 ttgaaaaatga ttgtctagct tgatagctct caccactgag gaaatgttct ctggcaaaaa  
 1500  
 cggcttctct cccaggaac tctgagaaag tgttattaag aaatgtggct tctactttct  
 1560  
 65 ctgtcttacg gggctaacat gccactcagt aatataataa tcgtggcagt ggtgactact  
 1620

ctcgtaatgt tgggtgcttat aatgtttctca tctctctcat tttccag  
1667

5 <210> 64  
<211> 195  
<212> ADN  
<213> Homo sapiens

10 <400> 64  
gtaactgcct tgagggagaa tggcacactt aagatagtagc cttctgctgg ctttctcagt 60  
gcacgagtat tgttcctttc cctttgaatt gtctattgc attctcattt gtagagtga  
120  
ggtttgttgc agatggggaa ggtttgtttt gttgtaaata aaataaagta tgggattctt  
15 180  
tccttgtgcc tttag  
195

20 <210> 65  
<211> 284  
<212> ADN  
<213> Homo sapiens

25 <400> 65  
gtctgttagg gcaagatcaa acagtgtcct actgtttgaa tgtgaaattc tctctcatgc 60  
tctccctgt tttctttgga tggcctttan ccaaggtagt agatccctac agagtccaaa  
120  
gagaagttag gaaatggta aagccacttg ttttttgcag catcngcat gtnatcaaac  
30 180  
ctganagagc ctatccatat cacttttctt taanagacat taaanatggn tccttaattc  
240  
cttttgancc cattgtattt attattcttt ttctgctggg gtcc  
284  
35

<210> 66  
<211> 560  
<212> ADN  
40 <213> Homo sapiens

<400> 66  
tctagaaaat ttttaggaac agaaaacttt ccagttctct caccctgct caaagagtgt 60  
atggctctta cattatat aactgcctga ctctacacag tatcagtact tagatcattt  
45 120  
gaaatgtgtc cacgttttac caaaatataa tagggtaga agctgagatg ctaattgcca  
180  
ttgtgtattc tcaaatatgt caagctacgt acatggcctg tttcatagag tagtctataa  
240  
50 gaaattgatg acttgattca tccgaatggc tggtgtgaac acctgggttac gcatgaacac  
300  
ctcttttcag ttgtctcaag acacctttct tttctgtact tatcagacaa ggactgaaag  
360  
gcagagactg ctactgttag acattttgag tcaagctttt ccttggacat agctttgtca  
55 420  
tgaaagccct ttacttctga gaaacttcta gcttcagaca catgccttca agatagttgt  
480  
tgaagacacc agaagaagga gcatggcaat gccgaaaaca cctaagataa taggtgacct  
540  
60 tcagtgttgg cttcttgcag  
560

<210> 67  
65 <211> 1649  
<212> ADN  
<213> Homo sapiens

<400> 67  
 gtgagacgtg ctgttttcgc cagagactct ggcttcatgg gtgggctgca ggctctgtga 60  
 ccagtgaagg caggatagca tcctgggtcaa gatatggatg ccggagccag atttatctgt  
 120  
 5 atttcaatcc cagttctatt ccttgccagt tgtgtatccg ctggcaagtt acttctctat  
 180  
 gcctcaatct cctcatctgt aaaatgggga taataatatt acctgcaata cagggttggt  
 240  
 acgaaaaata aatgaatag gtgcttagaa tggggcctga cattagtaag tgcttagttt  
 10 300  
 tgtgtgtgta tatgttattt ttatttttga ggagaacata aaaaggacaa agtgtagaaa  
 360  
 aactgggttg gtgtattcag ctgtcataac atgagagttg ttatgcccag atgcacttga  
 420  
 15 catgtgaatt tattagaac atgatttttc tctgagttga tgtttaactc aaactgatag  
 480  
 aaaagatagg tcagaatata gttggccaac agagaagact tgtagacta ttgtctgcat  
 540  
 gtcagtgttt gcatgctaac ttgcttagtt agaaaggtta aattttttca ctctataaaa  
 20 600  
 tcaagaaata tagagaaaag gtctgcagag agtctttcat ttgatgatgt ggatattggt  
 660  
 aagagcggga gtttggagca tacagagctc aagttgaatc ctgactttgc tacttattgg  
 720  
 25 ctatatgacc ttgggcaagc tgcttagtct ctctgaccc cagttacctt tgtttgttga  
 780  
 tgatgaccat tgataacaca accataaata atgacaacat agagatagtt ctctattatag  
 840  
 tagttgttat acagaattat tcaactcaatg ttaattttct gcattgaaat ccagaaacat  
 30 900  
 tagaattggg ggcattattt gaatctttaa ggttataagg aatacatttc tcagcaataa  
 960  
 atggaaggag ttttgggtta acttataaag tatacccaag tcattttttt ttcagagaag  
 1020  
 35 atatggtaga aagtcttagg aggttgaaga aggaattgga tatttattct ttctgagact  
 1080  
 atcatgggag ataatgacta tggttgtcca tgattggagc cgttgctgta gagttgggtc  
 1140  
 tattatagtg taggatttga atgggcatg tgttctcaga cctcagatta aatgagaaa  
 40 1200  
 actgaggcca gtggggagcg tgacttcaca tgggtacact tgtgctagag acagaaccag  
 1260  
 gattcaggac ttctggctcc tggctcctggg ttcatggccc aatgtagtct ttctcagtct  
 1320  
 45 tcaggaggag gaagggcagg acccagtgtt ctgagtcacc ctgaatgtga gcactattta  
 1380  
 ctctgtgaac ttcttggtt agtgccctctg ccagggtggcc ataacctctg gccttggtgt  
 1440  
 gccagagaaa aggttttagtt ttcaggctcc attgcttccc agctgccaag aatgccttgg  
 50 1500  
 tgcagcacag tcataggccc tgcattcttc attgccgtgc tggttgggtc gggaggtggg  
 1560  
 ctggactcgt agggatttgc cccttggcct tgtttctaac acttgccgtt tcctgctgtc  
 1620  
 55 ccctgcccc ctccactgcc tgggtaaaag  
 1649

<210> 68  
 60 <211> 1230  
 <212> ADN  
 <213> Homo sapiens

<400> 68  
 65 gtatgtttgt cttctacatc ccaggagggg gtaagattcg agcagaccaa agatgtttac 60  
 gagggccaag ggaatggact tcagaattac acggtggaat gaattttact gctgcggctc  
 120

aggtccctgt ataagctaata actgcatgca tagaacagca gcgaactaac cctgaataat  
 180  
 aggccagtct tctgttgagc ctttcagcct ctctcctctt catcctaactg ttgtcaggaa  
 240  
 5 cagccacatg tgttttaggt gaaataatcc acccttgcaa aaatccatga ttaagttata  
 300  
 aaatatttgg atttgtggag ctgtgtttta attctgtaac tgagtcacag ggcacactgt  
 360  
 10 caaagcatag aacctccaga gacttgtttt ctgcaaagta taattcatgt aattattatc  
 420  
 tattctgtta tatttgggat gttaggtagt gtttgttctt tagataaaaa tatccccac  
 480  
 tctgtaacaa tacattaaat caaagaaaag gacaaaggat ttttctgggt cttgttagca  
 540  
 15 ggagctttct tcagtcctga aagatttgta gacctgtaga tgggggaact gtgtcagtga  
 600  
 tacaaaaggg aagcatttaa aaaaaaaag tatatatata tatatatata tatgtaatgt  
 660  
 gaattggcct ctttttctct aagcccat tttctctta catagttcag gtttacttta  
 720  
 20 ttttttctt tccggctgct gacctgtat tgcccgtagt tgtggaacat agcatgtgtt  
 780  
 tgtgacctgt gcctgttatt tttgtgcttt ctagtgtgc atgcaaagag tacaaagt  
 840  
 25 tcttgccctt tcttgaaaa tctgtctgt ctgtgcaaa gggataattg tgaaagcact  
 900  
 tttgaaatac ttaatgagtt gattttcttc aaattaaaaa aaatatataa atgtatctgt  
 960  
 gtatgtacat gtgtgtacac atacacacct ttatacatag agccattta aaacaagctc  
 1020  
 30 cactttggag tgctctacgt caccctgatg ccgaatacag ggccagagtc tgagatcctt  
 1080  
 ctgggtggtt tctgtgtttt gttcatttct gttttaagag cctgtcacag agaaatgctt  
 1140  
 35 cctaaaatgt ttaatttata aaaacatttt tatctctcga ttactgggtt taatgaatta  
 1200  
 ctaagctggc tgccctctcat gtaccacag  
 1230  
 40  
 <210> 69  
 <211> 3035  
 <212> ADN  
 <213> Homo sapiens  
 45  
 <400> 69  
 gtgagtgccca ctttagccat aagcaggctt cttgtgcttg ttgcctgggt tgatttctaa 60  
 tatgctgcat ttatcaactg catgccacat tgtgaccgcc agcatttgcc ctttgaatta  
 120  
 50 ttattatggt ttatttaca aaagcgaagg tagtaaccga actaaattat ctaggaacaa  
 180  
 acgttttgag agtcttctaa caccgtgcaa agcacgtcat tacagacatt tgtttactga  
 240  
 tttagaacct taatatttaa tttaaatagc actttacact tactgatgaa atgcttttcc  
 300  
 55 tttctttctc tcccagcccc tgtacttaag tgcttcaata ggctctcatt atatatgatt  
 360  
 tttaggtttt gcttatcagc ttcttcgctt ttataatctg aaaagatggc atatgaattt  
 420  
 60 ttataaaaag ggacacttct ttcttctcaa attgtatatt tttattgtac tttccttcaa  
 480  
 aaccccttt taaaaagtaa gcagtggata aataaattca gtgaagcatc catatgacct  
 540  
 ttaagtgagt gtagggaag ggaggtcacc agatcactgt gagtgaagat ggtggagagg  
 600  
 65 tgaggatctt atgaggccgt gctcaaggct ggttagagggt ggttagtggt tccaggttta  
 660

ggcagaatct cagctgaggt catgaaacaa cagtgatctc tgaaaaatta tggcaagggt  
 720  
 ggaagggtgct ggagaattgg agagggggca aacttgactt tcaagtttca atgggaagat  
 780  
 5 aggtgactct gcacaccaca gaacagtgag catgataacc tgtttataca aggttctaga  
 840  
 gcagatttct aaatggatag ctactgtgtg cttgtttgtt cttaattagt attggatagt  
 900  
 tactaaatac ttgttagtac ttagtacata atgggtggt aatcctagca gctaatttg  
 10 960  
 gttcccaaat aaccagatga caaggataga gaaggacaca gacacggcct atctggattt  
 1020  
 catgtgcct ttcattttcc acatgaagggt tgtgtaggga agatagaagc atgagatgag  
 1080  
 15 atgataatat agttatctgg attcatcact ggcagctga accatatgaa ctcatggatt  
 1140  
 gatgctagct taggaaggct ctgtaggagc cagaactggg ctgagagcca gcccatagag  
 1200  
 acaaaagagg cccggccctg acatcagagg gtccaacat gatgtctgag ccccacctac  
 20 1260  
 agtctgccgg aggtgggtgg aaggaagagc ctttatcctt acaattctta ctgaaattca  
 1320  
 aatttttagg ttttgcaaaa aaatgggtga cctgaaggaa atttgacagg agcatgtctc  
 1380  
 25 agctgtattt aaatttgtct cagccaatcc ctttttgaat gttcagagtg taagcttcag  
 1440  
 gagggcagcg cgtcttagtg tgacttttct ggtcagttca ggtgctttta ggagacaatt  
 1500  
 agagatcaat ctggaaaact tcatttgaat ttttaataca taagaaaaca ataagaaata  
 30 1560  
 gttaaaaata tatatttata atatatatat gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt  
 1620  
 atatatatat attttattta tttatttttt tttagatgg agtctcgctc tgttggccag  
 1680  
 35 gctggagtgc agtggctcaa tcttggtcga ctgccacctc tgcctcccag gttcaagtga  
 1740  
 ttctcctacc tcagcctcct gagtagctgg gattacaagc atgtgccacc aactggcta  
 1800  
 atttttctaa ttttagtaga gatggagttt caccatgttg gacaggatgg tcttgaactc  
 40 1860  
 ctgacttagt gatccaccg ccttcgcctc ccaaagttct gggattacag gcatgagcca  
 1920  
 tcgtgcctgg caattatatt taatatttaa taataaggaa ataattgctg taactttact  
 1980  
 45 ttaaatgtg gaattctgaa actggaaggg aactggaaat gacttggtga atcaaatcat  
 2040  
 tttaaacttt tattttgcca gtggaaaaa taagcccca aaagagcagg ggacctgctg  
 2100  
 atgtcccaca gtaattcaga gctggagatg aggttgaagg ctttgtgtct tatctccagg  
 50 2160  
 gaaaatttgt agacagcgta gctctttatg tgacgagcat tctcaccoca gtcaccccc  
 2220  
 aattctctac tcatttgaga acataaattg gatcttgcca gtctctactc atttttcagc  
 2280  
 55 acatcgagca taagatccag actctttccc aggcctctct catctggctc ctctcctcct  
 2340  
 cttttatcat tactcttctt cgtagcttat cctactccag ccatgctgtc ttctattat  
 2400  
 tcctaaaaag tagaaatgca tttcttccca gggcctttgt acctgcaact gccatcgctt  
 60 2460  
 ttgctcagaa tgttcttttt gccaaagctt tgcccagctt gttctccatc attgttatgt  
 2520  
 tttggctgaa atgtcttctc ttagtaggtt cattctcccc agtcaactgtc tttttatttt  
 2580  
 65 gctttatttt gggccatcta aggttatctt attagtgtat ttgttggtcg tctcctccat  
 2640  
 gggcatacac ctccatgaag gcaggatatt tcaccttagg ccctcgaata tactggacag  
 2700

catctggcac gtagtagatg ctcaacgaat gtttgttgtg tgagcaaatg gttggttgat  
 2760  
 tggattgaac tgagttcagt atgtaatat ttagggcctc tttgcattct attttactta  
 2820  
 5 tgtataaaat gatacataat gatgatataa atgatgtcac agtgtacaag gctgttgtgg  
 2880  
 gatcaagcaa tcaaatgaga tcatgcttgt cttttccaaa tggtgaggga atagatgcat  
 2940  
 10 gtttgtggtt gttacggaat gatcctgtgc tcctgaggca acagaaaggc caggccatct  
 3000  
 ctggtaatcc tactcttgct gtcttccctt tgcag  
 3035

15 <210> 70  
 <211> 342  
 <212> ADN  
 <213> Homo sapiens

20 <400> 70  
 gtgagtcact ttcagggggt gattgggcag aaggggtgca ggatgggctg gtagcttccg 60  
 cttggaagca ggaatgagt agatatcatg ttgggagggt ctgtttcagt cttttttgtt  
 120  
 ttttgttttt ttttctgagg cggagtcttg ctctgtcgcc caggctggag tgctgtggca  
 180  
 25 tgatcttgcc tcaactgcaac ctccacctcc caggttcaag cgattctcct gcctcagcct  
 240  
 cctgagtagc tgggattaca ggcacgcacc accatgtctg gctaattttt gtgttttttag  
 300  
 30 tagagatagg gtttcgccgt gttggctagg ctggtctgga at  
 342

35 <210> 71  
 <211> 1182  
 <212> ADN  
 <213> Homo sapiens

40 <400> 71  
 gaattcctga cctcaggtga tccaccgcc tggcctccc aaagtgtctg gattacaggc 60  
 gtgagccact acgccagcc ctgtttcagt ctttaactcg cttcttgtca taagaaaaag  
 120  
 catgtgagtt ttgaggggag aaggttttga ccacactgtg cccatgcctg tcccacagca  
 180  
 45 gtaaaagtcac aggacagact gtggcaggcc tggcttcaa tcttggtctt gcaacaaatg  
 240  
 agctggtagc ctttgacagg cctgggcctg tttcttcacc tctgaattag ggaggctgga  
 300  
 ccagaaaact cctgtggatc ttgtcaactc tggattctt agagactctg tttgggaagg  
 360  
 50 agtcctgagc cttttttttt ttcttgagaa ttccaggaag aggagtgtt atgatagctc  
 420  
 tctgctgctt ttatcagcaa ccaaattgca ggatgaggac aagcaattct aatgagtagc  
 480  
 55 aggaactaaa agaaggcttg gttaccactc ttgaaaataa tagctagtcc aggtgcgggg  
 540  
 tggctcacac ctgtaatctc agtatttttg gatgccgagg tggactgac acctaaaggtc  
 600  
 aggagttcga aaccagcttg gccaatgttg cgaaaccctg tctctactaa aaattcaaaa  
 660  
 60 attagccagg catggtggca catgcctgta atcccagtta cttgggaggc tgaagcagga  
 720  
 gaattgcttg aacctgggag gtggaggtcg caggagacca aaattgcgcc actgtactcc  
 780  
 65 agcctgagca acacagcaaa actccatata aaaaaataaa atgaataaaa taacagctaa  
 840  
 tctagtcac agtataactc cagtgaacag aagatttatt aggcatagtg aatgatggtg  
 900

cttcctaaaa atctcttgac tacaagaat ctcatttcaa tgtttattgt ttagatgttc  
 960  
 agaataaatt cttgggaaag accttggtt ggtgtaagt aattaccagt gccgagggca  
 1020  
 5 ggggtgaacca agtctcagt ctggttgact gagggcagt tctgggacct gtagtcaggt  
 1080  
 ttccgggtcac actgtggaca tggtcactgt tgtccttgat ttgttttctg tttcaattct  
 1140  
 tgtctataaa gaccctgatg cttgggtttc atgtgatgac ag  
 10 1182

<210> 72  
 <211> 1309  
 15 <212> ADN  
 <213> Homo sapiens

<400> 72  
 gtgacttttt actaaacttg gccctgccg tattattact aattagagga attaaagacc 60  
 20 tacaataaac agactgaaac agtgggggaa atgccagatt atggcctgat tctgtctatt  
 120  
 ggaagtttag gatattatcc caaactagaa aagatgacga gagggactgt gaacattcag  
 180  
 ttgtcagctt caaggctgag gcagcctggt ctagaatgaa aatagaaatg gattcaacgt  
 25 240  
 caaattttgc cacttagtag caacttgacc aggttaactgg ttatcctttt aaagccttag  
 300  
 tttatctaaa ttgtgatatt aatgttgctc ttataagttt gtcattgagga cttaaattaaa  
 360  
 30 tgggtgtacat agagtgcctt ggggtactctc tgatggggga ctccatgata atttgtggtc  
 420  
 tcatggaggg agctctggga aggttttagga gcctgccttg gctctgcagc cttgggagag  
 480  
 ccttctagct tcccaggaca tggcagccta gtgttgaaatg cttggctcag caaatgtttg  
 35 540  
 ttctcgtttc cttcccatca acttggtcag ttggggctctt tcagtttagga gtatctcagt  
 600  
 gacttttaaat ggcattgggca tgctggagt atagtaccca tgagtttcta agaaagaagc  
 660  
 40 ataatttctc catatgtcat ccacaattga aatattattg ttaattgaaa aagcttctag  
 720  
 gccaggcacg gtggctcatg cctgtaatcc cagcacttta ggaggccaag gcgggtggat  
 780  
 cacttgaggt caggagtttg agaccagcct ggccaacatg gggaaaccct gtctctacta  
 45 840  
 aaaatataaaa ataagctggg cgtggtggtg cgtgcctgta atcccagcta cttgggaggc  
 900  
 tgaggcagga gaattgcttg aatctgggag gcggagggtg cagttagctg agttcatgcc  
 960  
 50 attgcattcc agcctgggca acaagagcga aaccatctcc caaaagaaaa aaaaaagaaa  
 1020  
 gaaaaagctt ctagtttggt tacatcttgg tctataaggt ggtttgtaaa ttggtttaac  
 1080  
 ccaaggcctg gttctcatat aagtaatagg gtatttatga tggagagaag gctggaagag  
 55 1140  
 gcctgaacac aggcctcttt tctctagcac aacctacaa ggccagctga ttctagggtt  
 1200  
 atttctgtcc gttccttata tctcaggtg gatatttact ctttttgcatt cattaggaat  
 1260  
 60 aggcctcagt ctttctttga actgattttt tgtttctttg tctctgcag  
 1309

<210> 73  
 65 <211> 1124  
 <212> ADN  
 <213> Homo sapiens



<400> 73  
 gtaagttgct gtctttctgg cacgtttagc tcagggggag gatgggtgtg taggtgtctt 60  
 ggattgaaga aagccttggg gattgtttgt cactcacaca ctgtgggtg ccattctcact  
 120  
 5 gtgaggagga cagaagccct gtgaacatgt ggagcacaca ggggcacaga cagatttaga  
 180  
 ttaggcctgc tttatagagt ttctgcctag agcatcatgg ctcatgccc agcagccct  
 240  
 ccagaggcct ctgaaatatt tgatatactg atttccttga ggagaatcag aaatctctg  
 300  
 10 cagggtgtcta gggatttcaa gtaagtagtg ttgtgagggg aatacctact tgtactttcc  
 360  
 ccccaacca gattcccgag gcttcttaag gactcaagga caatttctag gcatttagca  
 420  
 15 cgggactaaa aaggtcttag aggaaataag aagcgccaaa accatctctt tgcactgtat  
 480  
 ttcaacccat ttgtccttct gggttttgaa ggaacaggtg ggactgggga cagaagagtt  
 540  
 cttgaagcca gtttgtccat catggaaaat gagataggtg atgtggctac gtcagggggc  
 600  
 20 ccgaaggctc cttgttactg atttccgtct ttctctctg ccttttcccc aagggccag  
 660  
 acccctggat ctctggcag agcagacgca ggccccata atagccctca tgctagaaag  
 720  
 25 gagccggagc ctgtgtataa ggccagcgca gcctactctg gacagtgcag ggttccact  
 780  
 ctcccaactc cccatctgct tgcctccaga cccacattca cacacgagcc actgggttg  
 840  
 aggagcatct gtgagatgaa acaccattct ttccatcatg tctcagctat ctaactgtgt  
 900  
 30 gtgtaatcag gccaggctct ccctgctggg cagaaacat gggagttaag agattgccaa  
 960  
 catthattag aggaagctga cgtgtaactt ctctgaggca aaatttagcc ctctttgaa  
 1020  
 35 caggaatttg actcagtga ccttgtacac actcgactg agtctgctgc tgatgatact  
 1080  
 gtgcacccca ctgtctgggt ttaaatgtca ggctgttctt ttag  
 1124  
 40  
 <210> 74  
 <211> 1472  
 <212> ADN  
 <213> Homo sapiens  
 45  
 <400> 74  
 gtaaaatatct atcgtaatat gtatcagaaa aatgggcatg tagctgctgg gatataaggag  
 60  
 tagttggcag gttaaacgga tcacctggca gtcattgtt ctgaatatgt tggcatacag  
 120  
 50 agccgtcttt ggcatttagc gatttgagcc agacaaaact gaattactta gttgtacgtt  
 180  
 taaaagtgtg ggtcaaaaac aaatccagag gccaggagct gtggtcatg cctgtaatcc  
 240  
 55 tagcactttg ggaggccgaa gcgggtggat cacttgaggt caggagtctg agaccagcct  
 300  
 ggcctacatg acaaaacccc gtatctacta aaaatacaaa aaaattagct gggcttgggtg  
 360  
 gcacacacct gtaatcccag ctacttggga ggctgaggca ggagaattgc ttgaaccctg  
 420  
 60 taggaagagg ttgtagttag ccaagatcgc accgttgac tccagcctgg gcaacaagag  
 480  
 caaaactcca tctcaaaaaa caaattaaat ccagagattt aaaagctctc agaggctggg  
 540  
 65 cgcggtggct tacacctgtt atcccagcat tttgggatgc cgaggcgggc aaagcacaag  
 600  
 gtcaggagtt tgagaccagc ctggccaaca tagtgaaacc ctgtctctgc taaaaacata  
 660

gaaaaaattag ccgggcatgg tggcgtgcgc ctgtaatccc agctactcgg gaggctgagg  
 720  
 tgagagaatt acttgaaccc gggaggcgga ggttgcaagt agcccagatt gcaccactgc  
 780  
 5 actccagcct gggcgacaga gcaagactcc atctcaaaaa aagctctcag aacaaccagg  
 840  
 tttaaaaatt tggtcagttg gtaataaac tgggtttcaa acatactttg ctgaacaat  
 900  
 10 cactgactaa ataggaaatg aatctttttt tttttttttt aagctggcaa gctggtctgt  
 960  
 aggacctgat aagtactcac ttcatttctc tgtgtctcag gtttccatt tttaggtgag  
 1020  
 aattaagggg ctctgataaa acagacccta ggattgtgga cagcagtgat agtcctagag  
 1080  
 15 tccacaagtc tgcttttgag tgatgggccc atgtatcttg cacatctgca ggcagagcgt  
 1140  
 ggttctggct cttcagatga tgccggtgga gcactttgag gagtcctcac cccaccgtga  
 1200  
 taaccagaca ttaaaatctt ggggctttgc atcccaggat ttctctgtga ttccttctag  
 1260  
 20 acttgtggca tcatggcagc atcactgctg tagatttcta gtcacttggg tctcaggagc  
 1320  
 cgtttattta atggcttcac atttaatttc agtgaacaag gtagtggcat tgctcttcac  
 1380  
 25 agggccgtcc tgttgtccac aggttccaga ttgactgttg ccccttatct atgtgaacag  
 1440  
 tcacaactga ggcaggtttc tgttgtttac ag  
 1472  
 30  
 <210> 75  
 <211> 292  
 <212> ADN  
 <213> Homo sapiens  
 35  
 <400> 75  
 gtaaacgct gtctttgttc tagtagcttt ttgatgaaca ataactcta tgtttcttgg 60  
 agtactttca actcatggta aagttggcag gggcattcac aacagaaaag agcaaactat  
 120  
 40 taactttacc agtgaggcag tacggtgtag tgtagtatt cagagaattt gctttgccac  
 180  
 cagacatacc aggtaacctt gactaagtta cttaacctat ctaaacctca gttccctcat  
 240  
 ctgtgaaatg gagacagtaa tcatagctat ttccaaactg ttgtgagaat tc  
 45 292  
 <210> 76  
 <211> 235  
 50 <212> ADN  
 <213> Homo sapiens  
 <400> 76  
 gaattcaatg agttaaggt ataaggctct caccacagcg cctgcccaca tagtcagtga 60  
 55 tcactatgtc ctgaacactg taattacttc gccatattct ctgatcatag tgttttgcct  
 120  
 tggtagtgta ctagaatttc tttctgaggt ttatgggcat ggttgggtgg tatgcacctg  
 180  
 cctgcaggag cccggttttg gggcattacc ttgtacctgg tatgttttct ttcag  
 60 235  
 <210> 77  
 <211> 240  
 65 <212> ADN  
 <213> Homo sapiens  
 <400> 77

gtaagtgtgg ctgtgtctgt atagatggag tggggcaagg gagaggggta tggagaaggg 60  
gagaaaaatg tgaatctcat ttaggggaa cagctgcaga gaccgttata ttatgataaa  
120  
tctggattga tccaggtctt gggcagaagt gataagtta cgaattggct gggtgggctt  
5 180  
cttgaactgc agaagagaaa atgacactga tatgtaaaaa tcgtaacatt tagtgaattc  
240

10 <210> 78  
<211> 988  
<212> ADN  
<213> Homo sapiens

15 <400> 78  
gaattcatat aaagtgagtt caaaaattgt taattaaatt ataatttaata tataagtgtt 60  
taatcagttt gatttgttta aaaaccactg ttttaaattt ggtggaatat gtttttatta  
120  
gcttgtatct ttaattccta aattaagctg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg  
20 180  
tgtgtgtgtg aagtttaaaag ccaggatgag ctagttaaag gtatgcagcc tttggagtca  
240  
tacagatctg ggtttgaatc tggctcttaa actttataga tgtatgatat taaatgaggg  
300  
25 agttcatgta aattgccaaag cccagcactc agcacagagt tgatatttca cacacattag  
360  
atacctttcc tgtatgtgga gcatggcagt tctgttttct gctttactcc tacaggatac  
420  
taatatagga cactaggatc tttataccaa gaccccatgt aatgggctta tgagaccatt  
30 480  
cttcttataa aaatctgaca gaatttttgt atgtgttaga tcaataggct gcatactgtt  
540  
atthttcaagt tgatttacag ccagaaatat taatttattt gagtagttac agagtaatat  
600  
35 tctgtctctc atttagtttt caagcccccac tagtcccttg tgtgtgaaaa tttacaactt  
660  
actgctctta caaggtcatg aacagtggac caaagtgaat gccattaacc actctgactt  
720  
ccttcattag ttttattgtg acagtggact cttttgacct cagtaatacc agtttggcat  
40 780  
ttacattgtc atatttttag acttaaaaaat gatcatctta accctgaata aaatgtgtct  
840  
gggtgaacaga tgttttttct tgggtgtgtg ctcagatata tctgtgtgtg gtgtacgtgt  
900  
45 gtgtttgtct gtgtgtccat gtcctcactg attgagccct agctgcatca aaagaccctt  
960  
cagattttca cacgcttttt ctctccag  
988

50 <210> 79  
<211> 498  
<212> ADN  
<213> Homo sapiens

55 <400> 79  
gtaaggacac aggcctgctg tatctttctg atgtctgtca gggccatgga ttgatatgga 60  
taagaaagaa agagctcttg ctatcatcag gaaatgttcc agctactcta aagatgtatg  
120  
60 aaaaagaaat agccagagggc aggtgatcac tttcatgaca ccaaacacag cattgggtac  
180  
cagagttcat gtcacaccag agggaaaatt ctgtacacaa tgatgaaaat taataccact  
240  
accacttaag ttcctatgtg acaactttcc caagaatcag agagatacaa gtcaaaaactc  
300  
65 caagtcaatg cctctaactt ctctgatggg ttttaacctc cagagtcaga atgttctttg  
360

ccttactagg aaagccatct gtcatttgaa aactctgtac attttatcag cagcttatcc  
420  
atccattgca aatatgtttt tgtgccagcc acaatatatt gcttctatct ggaccaatag  
480  
5 ggggatttga aggaattc  
498

<210> 80  
10 <211> 544  
<212> ADN  
<213> Homo sapiens

<400> 80  
15 gaattctcat aattgtccta tcgtcaagtc tttatttctg cattttactg cttgatacac 60  
tgtcaggaca gactttaaaa ttattctcag tgcgatgaaa caattctgac attcatgtta  
120  
tgagcagtta cctcataaat agattacatg tgagattgaa cttgggcaga ctataatata  
180  
20 gcattaatga cgaaacagac acagtcacatc tcgggaagaa gaatagaggc ttatttgctg  
240  
cctgtgaaat taaaattact ctgactggga atccatcggt cagtaagttt actgagtgtg  
300  
acaccttggc ttgactgttg gaaagacaga aagggcattgt agtttataaa atcagccaag  
25 360  
gggaaaatgc ttgtcaaaat gtattgtcgg gtattttgat taatagttaa tgtggcttca  
420  
ttaattcaga gttactctcc aatatgttta tctgcccttt cttgtctgat aatggtgaaa  
480  
30 acttgtgtga tgcattgtat atttgattta ggggtgaact ggatgtcttt gttttcactt  
540  
ttag  
544

35 <210> 81  
<211> 111  
<212> ADN  
<213> Homo sapiens

40 <400> 81  
gtaagtcacc tctgagttag ggagctgcac agtggataag gcatttggtg cccagtgta 60  
gaaggagggc agggactctc agtagacact tatctttttg tgtctcaaca g  
111

45 <210> 82  
<211> 363  
<212> ADN  
50 <213> Homo sapiens

<400> 82  
gtgagtcatt cagagagaac actcctgctg ggatgagcat ctctgggagc cagaggacag 60  
55 tgttttaatt tgatcttatt ccacttgtca gtggtattga cactgctgac tgccttgtcc  
120  
tgtcttcaga gtctgtcttc cctgagaagg caaagcacct ttctttcttg ctgtgcctta  
180  
cattttgctg gtcaagcctt tcagtttctt ttgacagttt tttttacttc tttctttttt  
240  
60 caatgttgct cttaccaaga gtagctcttc tgccttccac ttacacatg agagctgggc  
300  
gacgccattc agtectaagg cttttaccat cacctctctt ggtgttttta ttgtcatctc  
360  
taa  
65 363

<210> 83

<211> 434  
<212> ADN  
<213> Homo sapiens

5 <400> 83  
tttaattgat tcactaggat atatgctact gaaaggggaa tctgcttaaa gtgctttctg 60  
atatttatta ttactaaaac ttagaattta ttaaaaatac tgactgtgaa aaattacttg  
120  
ggctcgtttgc ctttttaaaa ggatttttgg catgtctcat taaaaaaga aatactagat  
10 180  
atcttcagtg aagttacaaa tcgaatacac attggctctg aaattctgat tgatactggg  
240  
tcataaaaag ttttcccaaa tcagacttgg aaagtgatca ctctcttggt actctttttt  
300  
15 ccttgcctatg ggtgatagcc atttgtgttt attggaagat cgggaattt taaggaacat  
360  
aggcccaaat ttgaggaagg gccatggttt ttgatccctc cattctgacc ggatctctgc  
420  
attgtgtcta ctag  
20 434

<210> 84  
<211> 264  
25 <212> ADN  
<213> Homo sapiens

<400> 84  
gtgagctttt tcttagaacc cgtggagcac ctggttgagg gtcacagagg aggcgcacag 60  
30 ggaaacactc accaatgggg gtgcatga actgaactca aaatatgtga taaaactgat  
120  
tttccctgatg tgggcacccc gcagccccct cctgccccat cctggagact gtggcaagta  
180  
ggttttataa tactacgtta gagactgaat ctttgcctg aaaaatagtt tgaaagggtc  
35 240  
atttttcttg ttttttcccc caag  
264

40 <210> 85  
<211> 175  
<212> ADN  
<213> Homo sapiens

45 <400> 85  
gtgagagagt acagggttaca atagctcatc ttcagttttt ttcagcttta tgtgctgtaa 60  
cccagcagtt tgctgacttg cttaataaaa gggcatgtgt tcccaaatg tacatctata  
120  
ccaaggttct gtcaatttta ttttaaaaac accatggaga cttcttaaag aattc  
50 175

<210> 86  
<211> 588  
55 <212> ADN  
<213> Homo sapiens

<400> 86  
gaattcccat tctcgaatac attggtttta tatgcttaca tttatgtgtt agttattaaa 60  
60 acatactaatt attgtatatc tagtcaaaac tgaggtagag agaataaatg gttgattttg  
120  
agtttgagtt tcatagtcca aaaagctgat atattgcctg tgttcaagag ggtctatata  
180  
agccctctag atgccagcat ctccaaattt tacttttttg gaatctgtac agtatttgca  
65 240  
atatttttat tacaaatttc tactctgtgg aatttaattt ttaaaatacc tgcaatacat  
300

atatatgttg aatagatgaa aaattatgta gataataatg aatgatacgg ttctaaaaag  
 360  
 acagggttaaa aagtaagttc acttttattt tgagcttcag aatcattcag aagccagtcg  
 420  
 5 ccacaaacgc agaccaaggc tcttggcaca tcaaataatgc ctatggctta gggttattga  
 480  
 caagtcttat gttgcagtgat atgtgggtta tagtctgcc ttccacagtt gcttgggaga  
 540  
 10 gctgtgagtc actgaggctt atgaatgttt acattttgtt tgttgag  
 588

<210> 87  
 <211> 78  
 15 <212> ADN  
 <213> Homo sapiens

<400> 87  
 gtaaagacac tttgtctata ttgcgtttgt ccttattagt tcagactatc tctacccaat 60  
 20 caagcaacga tgctcgtt 78

<210> 88  
 <211> 376  
 25 <212> ADN  
 <213> Homo sapiens

<400> 88  
 acatgtgccg gtactgggtga gagcgcaagc ttggagtgca aacacaaatg gggttgcac 60  
 30 ctggccctac caattatgag ctctgagcca tgggcaagtg actaactccc tgggcctcag  
 120  
 tttctctgta acatctgtca gacttcatgg gtccagggtga ggattaaagg agatcatgta  
 180  
 ttacagcac atggcatggg gcttcacata aaataagtat ttagtaaatg ataactgggt  
 35 240  
 ccttctctca gaaacttatt tctgggcctg ccagggggccg ccttttttca tggcacaagt  
 300  
 tgggttccca gggttcagta ttcttttaaa tagttttctg gagatcctcc atttgggtat  
 360  
 40 tttttcctgc tttcag  
 376

<210> 89  
 45 <211> 111  
 <212> ADN  
 <213> Homo sapiens

<400> 89  
 50 gtaagctttg agtgtaaaaa cagatttact tctcagggtg tggattcctg ccccgacact 60  
 cccgcccata ggtccaagag cagtttgtat ctggaattgg tgcttgaatt c  
 111

55 <210> 90  
 <211> 2264  
 <212> ADN  
 <213> Homo sapiens

60 <400> 90  
 ctgtatagta atcattgact aaagccattt gtctgtgttt tcttcttggt gttgtatata 60  
 tcaggtaaaa ttttttccaa agagccatgt gtcagtgaat actgaaccac tttgatattg  
 120  
 agacattaat ttgtaccctg tgttattatc tactagtaat aatgtaatac tgtagaaata  
 180  
 65 ttgtctaat tcttttcaaa attgttgcac ccccttaga atgtttctat ttccataagg  
 240

atttaggtat gctattatcc cttcttatac cctaagatga agctgttttt gtgctctttg  
 300  
 ttcacatttg gccctcattc caagcacttt acgctgtctg taatgggac tatttttgca  
 360  
 5 ctggaatatac tgagaattgc aaaactagac aaaagtttca caacagattt tctaagttaa  
 420  
 atcattttca ttaaaaggaa aaaagaaaaa aaatttttgt atgtcaataa cctttatatg  
 480  
 10 aagtattaaa atgcattttt ctatgttgta atataatgag tcacaaaata aagctgtgac  
 540  
 agttctgttg gtctacagaa atttactttt gtgcatttgt ggcaccacct actgttgaag  
 600  
 gggtataaag ccattagaaa agtagagggg aagtgatttg gatcaaaagg aaaaacttta  
 660  
 15 gaaaagattc aaatgttccc ttaatcataa aagagaactg aggggactac ttgaaaataa  
 720  
 aagggtgttt tgtattttca tgttggttaa gatactgagt aactggtaact aagtgttaga  
 780  
 20 gggtttttaga taaatattct gcttaatgat tatgaagctg cactgagatt tctgaaaatg  
 840  
 ctctgtagct gagcttattt aataaatgtt cacttggtat aggggaagct acaaaggcag  
 900  
 ccttcagtgt ccttttgttt attcaaccaa aaatataagg acacaatgta gcagttatac  
 960  
 25 tgggaagggtg ctgggggttg tggcaatggt gagcaggaag gcgaagtaga tatggaaaca  
 1020  
 gaaatgatac taatatcggg gattccttcc ttttttctg taataagtgc tgtgcagaca  
 1080  
 acatatgagc agtgcgtgata aatgtaaata tatttttcat agctcattaa gaatcagttt  
 1140  
 30 cagaaaagaga tgtctgctta ttttgctact tgaagaatcc ctgtcaaaca gtccttttga  
 1200  
 ggaagtacaa gaggtgtctt ctatttgtga cctcaggaat ggctgtgaca gtgtcgtgag  
 1260  
 35 cagtcctttt cctgtggcac agatctgaac tttgtgtgca gaaaaatctt ggcttcaagt  
 1320  
 gagccaagat gccccctgag catcagcatc acaacttcac cctcctatct tgaagttcat  
 1380  
 40 gttatagtga ctttaatgaa atcatagaac actgtttctt cgtgaacaat gacgagggag  
 1440  
 aggaaaaaac tttattgaaa aataaaaagg caggtaattt agatgaaaat atgttaccca  
 1500  
 tgagggtttg tttttgcttt ttgttttgt ttttgagaaa cagaatctcg ctctgtcgtc  
 1560  
 45 caggctggag tgcagcggca tgatcttggc tcaactgcaac ctccgcctcc cgggttcaag  
 1620  
 cgattctcct cagcttccca agtagctggc actacaggca tgcgccacca caaccagcta  
 1680  
 50 atttttgtat ttttagtaga gatgggggtt cactatacgt tggccaggct ggtctcaaac  
 1740  
 tcctgacctt aggtgatcct tctgccttgg gctcccaaag tgctgggatt acaggcatga  
 1800  
 gccaccttgc ctggccctac ccattgacct tgactaaaac attcttctat ctgtagaaaa  
 1860  
 55 gcccaaaaga acttttccag attcaaaaaa cttggcactt tgtaatggta atgtttacat  
 1920  
 taagtaaaaa aaaaaaaaaa aaaccacatt agcttcagtt ttcaagtgtt tactgtgttg  
 1980  
 60 tcatgcactt catttaattc tcaacacctg ccctatgagg taaaaagtac cattttacat  
 2040  
 atgagtaaat tacagctcag tggataagaa actcgtccaa aggtacaggt tcagtcaagt  
 2100  
 ggcagagggt tctttttgtt gaagttaggt atcagttaaa attgaccttg taaaatcaca  
 2160  
 65 tcagcatcaa tatacattaa tttaacaaat atttattgaa ctttactgta tgccagatac  
 2220  
 ttctctaggt actagggggg acaatgtaga agaaaataga attc  
 2264

<210> 91  
 <211> 9497  
 5 <212> ADN  
 <213> Homo sapiens

<400> 91  
 caaacatgtc agctgttact ggaagtggcc tggcctctat ttatcttcct gatcctgac 60  
 10 tctgttcggc tgagctaccc accctatgaa caacatgaat gccattttcc aaataaagcc  
 120 atgcccctctg caggaacact tccttgggtt caggggatta tctgtaatgc caacaacccc  
 180 tgtttccgtt acccgactcc tggggaggct cccggagttg ttggaaactt taacaaatcc  
 15 240 attgtggctc gcctgttctc agatgctcgg aggcttcttt tatacagcca gaaagacacc  
 300 agcatgaagg acatgcgcaa agttctgaga acattacagc agatcaagaa atccagctca  
 360 aacttgaagc ttcaagattt cctgggtggac aatgaaacct tctctgggtt cctgtatcac  
 20 420 aacctctctc tcccaaagtc tactgtggac aagatgtga gggctgatgt cattctccac  
 480 aagggtatltt tgcaaggcta ccagttacat ttgacaagtc tgtgcaatgg atcaaaatca  
 25 540 gaagagatga ttcaacttgg tgaccaagaa gtttctgagc tttgtggcct accaaggagg  
 600 aaactggctg cagcagagcg agtacttcgt tccaacatgg acatcctgaa gccaatcctg  
 660 agaacactaa actctacatc tcccttcccg agcaaggagc tggccgaagc cacaaaaaca  
 30 720 ttgctgcata gtcttgggac tctgcccag gagctgttca gcatgagaag ctggagtgc  
 780 atgcgacagg aggtgatgtt tctgaccaat gtgaacagct ccagctcctc caccctaatc  
 35 840 taccaggctg tgtctcgtat tgtctcgagg catcccagg gaggggggct gaagatcaag  
 900 tctctcaact ggtatgagga caacaactac aaagccctct ttggaggcaa tggcactgag  
 960 gaagatgctg aaaccttcta tgacaactct acaactcctt actgcaatga tttgatgaag  
 40 1020 aatttgaggt ctagtctctt tcccgcatt atctggaaag ctctgaagcc gctgctcgtt  
 1080 gggaagatcc tgtatacacc tgacactcca gccacaaggc aggtcatggc tgagggtgaac  
 45 1140 aagaccttcc aggaactggc tgtgttccat gatctggaag gcatgtggga ggaactcagc  
 1200 cccaagatct ggaccttcat ggagaacagc caagaaatgg accttgtccg gatgctgttg  
 1260 gacagcaggg acaatgacca cttttgggaa cagcagttgg atggcttaga ttggacagcc  
 50 1320 caagacatcg tggcgttttt ggccaagcac ccagaggatg tccagtccag taatggttct  
 1380 gtgtacacct ggagagaagc tttcaacgag actaaccagg caatccggac catatctcgc  
 55 1440 ttcatggagt gtgtcaacct gaacaagcta gaaccatag caacagaagt ctggctcatc  
 1500 aacaagtcca tggagctgct ggatgagagg aagttctggg ctggtattgt gttcactgga  
 1560 attactccag gcagcattga gctgccccat catgtcaagt acaagatccg aatggacatt  
 60 1620 gacaatgtgg agaggacaaa taaaatcaag gatgggtact gggaccctgg tcctcgagct  
 1680 gaccctttg aggacatgcg gtacgtctgg gggggcttcg cctacttgca ggatgtgggtg  
 65 1740 gagcaggcaa tcatcagggt gctgacgggc accgagaaga aaactgggtg ctatatgcaa  
 1800



cagatgccct atccctgtta cgttgatgac atctttctgc gggatgatgag ccgggtcaatg  
 1860  
 cccctcttca tgacgctggc ctggatttac tcagtggctg tgatcatcaa gggcatcgtg  
 1920  
 5 tatgagaagg aggacaggct gaaagagacc atgcggatca tgggcctgga caacagcatc  
 1980  
 ctctggttta gctggttcat tagtagcctc attcctcttc ttgtgagcgc tggcctgcta  
 2040  
 10 gtggtcatcc tgaagttagg aaacctgctg ccctacagtg atcccagcgt ggtgtttgtc  
 2100  
 ttctgtgccg tgtttgctgt ggtgacaatc ctgcagtgtc tcctgattag cacactcttc  
 2160  
 tccagagcca acctggcagc agcctgtggg ggcacatctc acttcacgct gtacctgccc  
 2220  
 15 tacgtccctg gtgtggcatg gcaggactac gtgggcttca cactcaagat cttecgctagc  
 2280  
 ctgctgtctc ctgtggcttt tgggtttggc tgtgagtact ttgccctttt tgaggagcag  
 2340  
 20 ggcattggag tgcagtggga caacctgttt gagagtccctg tggaggaaga tggcttcaat  
 2400  
 ctacaccatt cgggtctccat gatgctgttt gacaccttcc tctatggggt gatgacctgg  
 2460  
 tacattgagg ctgtctttcc aggccagtac ggaattccca ggccttggtt tttccttgc  
 2520  
 25 accaagtcct actggtttgg cgaggaaagt gatgagaaga gccacctgg ttccaaccag  
 2580  
 aagagaatat cagaaatctg catggaggag gaaccacccc acttgaagct gggcgtgtcc  
 2640  
 attcagaacc tggtaaaagt ctaccgagat gggatgaagg tggctgtcga tggcctggca  
 2700  
 30 ctgaattttt atgagggcca gatcacctcc ttcctgggcc acaatggagc ggggaagacg  
 2760  
 accaccatgt caatcctgac cgggttgttc ccccgacct cgggcaccgc ctacatcctg  
 2820  
 35 ggaagagaca ttcgctctga gatgagcacc atccggcaga acctgggggt ctgtccccag  
 2880  
 cataacgtgc tgtttgacat gctgactgtc gaagaacaca tctggttcta tggccgcttg  
 2940  
 aaagggtctc ctgagaagca cgtgaaggcg gagatggagc agatggccct ggtgtttgtt  
 3000  
 40 ttgccatcaa gcaagctgaa aagcaaaaaca agccagctgt caggtggaat gcagagaaaq  
 3060  
 ctatctgtgg ccttgccctt tgtcggggga tctaagggtt tcattctgga tgaaccaca  
 3120  
 45 gctggtgtgg acccttactc ccgcagggga atatgggagc tgctgctgaa ataccgacaa  
 3180  
 ggccgcacca ttattctctc tacacaccac atggatgaag cggacgtcct gggggacagg  
 3240  
 attgccatca tctcccatgg gaagctgtgc tgtgtgggct cctccctgtt tctgaagaac  
 3300  
 50 cagctgggaa caggctacta cctgacctg gtcaagaaag atgtggaatc ctccctcagt  
 3360  
 tctgcagaa acagtagtag cactgtgtca tacctgaaaa aggaggacag tgtttctcag  
 3420  
 55 agcagttctg atgtggcct gggcagcgac catgagagtg acacgctgac catcgatgtc  
 3480  
 tctgctatct ccaacctcat caggaagcat gtgtctgaag cccggctggg ggaagacata  
 3540  
 gggcatgagc tgacctatgt gctgccatat gaagctgcta aggagggagc ctttgtggaa  
 3600  
 60 ctctttcatg agattgatga ccgctctca gacctgggca tttctagtta tggcatctca  
 3660  
 gagacgaccc tggagaaaat attcctcaag gtggccgaag agagtggggg ggtgctgag  
 3720  
 65 acctcagatg gtaccttgcc agcaagacga aacaggcggg ccttcgggga caagcagagc  
 3780  
 tgtcttcgcc cgttcaactga agatgatgct gctgatccaa atgattctga catagacca  
 3840

gaatccagag agacagactt gctcagtggg atggatggca aagggtccta ccaggtgaaa  
 3900  
 ggctggaaac ttacacagca acagtttgtg gcccttttgt ggaagagact gctaattgcc  
 3960  
 5 agacggagtc ggaaaggatt ttttgctcag attgtcttgc cagctgtgtt tgtctgcatt  
 4020  
 gcccttgtgt tcagcctgat cgtgccaccc ttgggcaagt accccagcct ggaacttcag  
 4080  
 10 ccctggatgt acaacgaaca gtacacattt gtcagcaatg atgtcctga ggacacggga  
 4140  
 accctggaac tcttaaacgc cctcaccaaa gaccctggct tcgggacccg ctgtatggaa  
 4200  
 ggaaacccaa tcccagacac gccctgccag gcaggggagg aagagtggac cactgcccc  
 4260  
 15 gttccccaga ccatcatgga cctcttccag aatgggaact ggacaatgca gaacccttca  
 4320  
 cctgcatgcc agtgtagcag cgacaaaatc aagaagatgc tgcctgtgtg tccccaggg  
 4380  
 gcaggggggc tgcctcctcc acaaagaaaa caaacactg cagatatacct tcaggacctg  
 4440  
 20 acaggaagaa acatttcgga ttatctggtg aagacgtatg tgcagatcat agccaaaagc  
 4500  
 ttaaagaaca agatctgggt gaatgagttt aggtatggcg gcttttccct ggggtgcagt  
 4560  
 25 aatactcaag cacttcctcc gagtcaagaa gttaatatgag ccaccaaaca aatgaagaaa  
 4620  
 cacctaaagc tggccaagga cagttctgca gatcgatttc tcaacagctt gggaagattt  
 4680  
 atgacaggac tggacaccag aaataatgtc aagggtgtgt tcaataacaa gggctggcat  
 4740  
 30 gcaatcagct ctttctgaa tgtcatcaac aatgccattc tccgggccaa cctgcaaaag  
 4800  
 ggagagaacc ctagccatta tgggaattact gctttcaatc atccctgaa tctaccaag  
 4860  
 35 cagcagctct cagaggtggc tccgatgacc acatcagtgg atgtccttgt gtccatctgt  
 4920  
 gtcattcttt caatgtcctt cgtcccagcc agctttgtcg tattcctgat ccaggagcgg  
 4980  
 gtcagcaaaag caaacacct gcagttcatc agtggagtga agcctgtcat ctactggctc  
 5040  
 40 tctaattttg tctgggatat gtgcaattac gttgtcccctg ccacactgggt cattatcatc  
 5100  
 ttcatctgct tccagcagaa gtcctatgtg tccctcacca atctgcctgt gctagccctt  
 5160  
 45 ctacttttgc tgtatgggtg gtcaatcaca cctctcatgt acccagcctc ctttgtgttc  
 5220  
 aagatcccca gcacagccta tgttgtgtct accagcgtga acctcttcat tggcattaat  
 5280  
 ggcagcgtgg ccacctttgt gctggagctg ttcaccgaca ataagctgaa taatatcaat  
 5340  
 50 gatatacctga agtccgtgtt cttgatcttc ccacattttt gcctgggacg agggctcatc  
 5400  
 gacatgggtga aaaaccaggc aatggctgat gccctggaaa ggtttgggga gaatcgcttt  
 5460  
 55 gtgtcaccat tatcttggga cttgggtgga cgaaacctct tcgccatggc cgtggaaggg  
 5520  
 gtggtgttct tctcattac tgttctgac cagtacagat tcttcatcag gccagacct  
 5580  
 gtaaatgcaa agctatctcc tctgaatgat gaagatgaag atgtgaggcg ggaaagacag  
 5640  
 60 agaattcttg atggtggagg ccagaatgac atcttagaaa tcaaggagtt gacgaagata  
 5700  
 tatagaagga agcggaagcc tgctgttgac aggatattcg tgggcattcc tctggtgag  
 5760  
 65 tgcttgggc tctgggagt taatggggct ggaaaatcat caactttcaa gatgttaaca  
 5820  
 ggagatacca ctgttaccag aggagatgct ttccttaaca gaaatagtat cttatcaaac  
 5880

atccatgaag tacatcagaa catgggctac tgcctcagt ttgatgccat cacagagctg  
 5940  
 ttgactggga gagaacacgt ggagttcttt gcccttttga gaggagtccc agagaaagaa  
 6000  
 5 gttggcaagg ttggtgagtg ggcgattcgg aaactgggcc tcgtgaagta tggagaaaaa  
 6060  
 catgctggta actatagtgg aggcaacaaa cgcaagctct ctacagccat ggctttgatc  
 6120  
 10 ggcgggcctc ctgtggtgtt tctggatgaa cccaccacag gcatggatcc caaagcccgg  
 6180  
 cggttcttgt ggaattgtgc cctaagtgtt gtcaaggagg ggagatcagt agtgcttaca  
 6240  
 tctcatagta tggagaagt tgaagctctt tgcactagga tggcaatcat ggtcaatgga  
 6300  
 15 aggttcagggt gccttggcag tgtccagcat ctaaaaata ggtttggaga tggttatata  
 6360  
 atagttgtac gaatagcagg gtccaacccg gacctgaagc ctgtccagga tttctttgga  
 6420  
 cttgcatttc ctggaagtgt tccaaaagag aaacaccgga acatgctaca ataccagctt  
 6480  
 20 ccatcttcat tatcttctct ggccaggata ttcagcatcc tctcccagag caaaaagcga  
 6540  
 ctccacatag aagactactc tgtttctcag acaacacttg accaagtatt tgtgaacttt  
 6600  
 25 gccaggacc aaagtgatga tgaccactta aaagacctct cattacacaa aaaccagaca  
 6660  
 gtatggacg ttgcagttct cacatctttc ctacaggatg agaaagtga agaaagctat  
 6720  
 gtatgaagaa tcctgttcat acgggggtggc tgaaagttaa gaggnactag actttccttt  
 6780  
 30 gcaccatgtg aagtgttgtg gagaaaagag ccagaagttg atgtgggaag aagtaaactg  
 6840  
 gatactgtac tgatactatt caatgcaatg caattcaatg caatgaaaac aaaattccat  
 6900  
 35 tacaggggca gtgcctttgt agcctatgtc ttgtatggct ctcaagtga agacttgaat  
 6960  
 ttagtttttt acctatacct atgtgaaact ctattatgga acccaatgga catatgggtt  
 7020  
 tgaactcaca cttttttttt ttttttgttc ctgtgtattc tcattggggt tgcaacaata  
 7080  
 40 attcatcaag taatcatggc cagcgattat tgatcaaaat caaaaggtaa tgcacatcct  
 7140  
 cattcactaa gccatgccat gcccaggaga ctggtttccc ggtgacacat ccattgctgg  
 7200  
 45 caatgagtgt gccagagtta ttagtgccaa gtttttcaga aagtttgaag caccatgggtg  
 7260  
 tgtcatgctc acttttgtga aagctgctct gctcagagtc tatcaacatt gaatatcagt  
 7320  
 tgacagaatg gtgccatgcg tggctaacaat cctgctttga ttccctctga taagctgttc  
 7380  
 50 tgggtggcagt aacatgcaac aaaaatgtgg gtgtctctag gcacgggaaa cttggttcca  
 7440  
 ttgttatatt gtcctatgct tcgagccatg ggtctacagg gtcacctta tgagactctt  
 7500  
 55 aaatatactt agatcctggt aagaggcaaa gaatcaacag ccaaactgct ggggctgcaa  
 7560  
 gctgctgaag ccagggcagt ggattaaaga gattgtgcgt tcaaacctag ggaagcctgt  
 7620  
 gcccatthgt cctgactgtc tgctaacatg gtacactgca tctcaagatg tttatctgac  
 7680  
 60 acaagtgtat tatthctggc tttttgaatt aatctagaaa atgaaaagat ggagttgtat  
 7740  
 tttgacaaaa atgtttgtac ttttaaatgt tatthggaat ttaagtctt atcagtgtat  
 7800  
 65 tctgaatcct tagaatggcc tctthgtaga accctgtggt atagaggagt atggccactg  
 7860  
 cccactatt tttatthct tatgtaagt tgcatactag tcatgactag tgcctagaaa  
 7920

gcaatgtgat ggtcaggatc tcatgacatt atatttgagt ttctttcaga tcatttagga  
7980  
tactcttaat ctcaattcat caatcaaata ttttttgagt gtatgctgta gctgaaagag  
8040  
5 tatgtacgta cgtataagac tagagagata ttaagtctca gtacacttcc tgtgccatgt  
8100  
tattcagctc actggtttac aaatataggt tgtcttggtg ttgtaggagc ccactgtaac  
8160  
10 aatactgggc agcctttttt ttttttttta attgcaacaa tgcaaaagcc aagaaagtat  
8220  
aagggtcaca agtctaaaca atgaattctt caacagggaa aacagctagc ttgaaaactt  
8280  
gctgaaaaac acaacttggt tttatggcat ttagtacctt caaataattg gctttgcaga  
8340  
15 tattggatac ccatttaa atgacagctct caaatttttc atctcttcaa tcaactagtca  
8400  
agaaaaatat aaaaacaaca aatacttcca tatggagcat ttttcagagt tttctaacc  
8460  
agtcttattt ttctagtcag taaacatttg taaaaaact gtttactaa tacttactgt  
20 8520  
taactgtctt gagagaaaag aaaaatatga gagaactatt gtttggggaa gttcaagtga  
8580  
tctttcaata tcattactaa cttcttccac tttttccaaa atttgaatat taacgctaaa  
8640  
25 ggtgtaagac ttcagatttc aaattaatct ttctatat tttaaattta cagaatatta  
8700  
tataaccac tgctgaaaaa gaaaaaatg attgttttag aagttaagt caatattgat  
8760  
tttaaatata agtaatgaag gcataatttc aataactagt gatatggcat cgttgcat  
30 8820  
tacagtatct tcaaaaatac agaatttata gaataatttc tctcattta atatttttca  
8880  
aatcaaatg tatggtttcc tcattttact aaaatcgtat tctaattctt cattatagta  
8940  
35 aatctatgag caactcctta cttcggttcc tctgatttca aggccatatt ttaaaaaatc  
9000  
aaaaggcact gtgaactatt ttgaagaaaa cacaacattt taatacagat tgaaggacc  
9060  
40 tcttctgaag ctagaaacaa tctatagtta tacatcttca ttaatactgt gttaccttt  
9120  
aaaatagtaa ttttttacat ttctctgtgt aaacctaatt gtggtagaaa tttttacca  
9180  
ctctatactc aatcaagcaa aatttctgta tttccctgtt ggaatgtacc tatgtgagtt  
9240  
45 tcagaaatc tcaaaatagc tgttcaaaaa tttctgcttt tgcattcttg ggacacctca  
9300  
gaaaacttat taacaactgt gaatatgaga aatacagaag aaaataataa gccctctata  
9360  
cataaatgcc cagcacaatt cattgtttaa aaacaaccaa acctcacact actgtatttc  
50 9420  
attatctgta ctgaaagcaa atgctttgtg actattaaat gttgcacatc attcattcaa  
9480  
aaaaaaaaa aaaaaaa  
9497  
55  
  
<210> 92  
<211> 2617  
<212> ADN  
60 <213> Homo sapiens  
  
<400> 92  
CAATGAAA CAAAATTCCA TTACAGGGGC AGTGCCTTTG TAGCCTATGT CTTGTATGGC 60  
TCTCAAGTGA AAGACTTGAA TTTAGTTTTT TACCTATACC TATGTGAAAC TCTATTATGG  
65 120  
AACCCAATGG ACATATGGGT TTGAACTCAC ACTTTTTTTT TTTTTTGTT CCTGTGTATT  
180

CTCATTGGGG TTGCAACAAT AATTCATCAA GIAATCATGG CCAGCGATTA TTGATCAAAA  
240  
TCAAAAAGGTA ATGCACATCC TCATTCACTA AGCCATGCCA TGCCCAGGAG ACTGGTTTCC  
300  
5 CGGTGACACA TCCATTGCTG GCAATGAGTG TGCCAGAGTT ATTAGTGCCA AGTTTTTCAG  
360  
AAAGTTTGAA GCACCATGGT GTGTCATGCT CACTTTTGTG AAAGCTGCTC TGCTCAGAGT  
420  
10 CTATCAACAT TGAATATCAG TTGACAGAAT GGTGCCATGC GTGGCTAACA TCCTGCTTTG  
480  
ATTCCCTCTG ATAAGCTGTT CTGGTGGCAG TAACATGCAA CAAAAATGTG GGTGTCTCTA  
540  
GGCACGGGAA ACTTGGTTCC ATTGTTATAT TGTCTATGC TTCGAGCCAT GGGTCTACAG  
600  
15 GGTCACTCTT ATGAGACTCT TAAATATACT TAGATCCTGG TAAGAGGCAA AGAATCAACA  
660  
GCCAAACTGC TGGGGCTGCA AGCTGCTGAA GCCAGGGCAT GGGATTAAAG AGATTGTGCG  
720  
20 TTCAAACCTA GGAAGCCTG TGCCCATTTG TCCTGACTGT CTGCTAACAT GGTACACTGC  
780  
ATCTCAAGAT GTTTATCTGA CACAAGTGTA TTATTCTTGG CTTTTTGAAT TAATCTAGAA  
840  
AATGAAAAGA TGGAGTTGTA TTTTGACAAA AATGTTTGTA CTTTTTAATG TTATTTGGAA  
900  
25 TTTTAAGTTC TATCAGTGAC TTCTGAATCC TTAGAATGGC CTCTTTGTAG AACCCTGTGG  
960  
TATAGAGGAG TATGGCCACT GCCCACTAT TTTTATTTTC TTATGTAAGT TTGCATATCA  
1020  
30 GTCATGACTA GTGCCTAGAA AGCAATGTGA TGGTCAGGAT CTCATGACAT TATATTTGAG  
1080  
TTTCTTTTCAG ATCATTTAGG ATACTCTTAA TCTCACTTCA TCAATCAAAT ATTTTTTGAG  
1140  
TGTATGCTGT AGCTGAAAGA GTATGTACGT ACGTATAAGA CTAGAGAGAT ATTAAGTCTC  
1200  
35 AGTACACTTC CTGTGCCATG TTATTGAGCT CACTGGTTTA CAAATATAGG TTGTCTTGTTG  
1260  
GTTGTAGGAG CCCACTGTAA CAATACTGGG CAGCCTTTTT TTTTTTTTTT AATTGCAACA  
1320  
40 ATGCAAAAGC CAAGAAAGTA TAAGGCTCAC AAGTCTAAAC AATGAATTCT TCAACAGGGA  
1380  
AAACAGCTAG CTTGAAACT TGCTGAAAAA CACAACCTGT GTTTATGGCA TTTAGTACCT  
1440  
TCAAATAATT GGCTTTGCAG ATATTGGATA CCCCATTAAT TCTGACAGTC TCAAATTTTT  
1500  
45 CATCTCTTCA ATCACTAGTC AAGAAAAATA TAAAAACAAC AAATACTTCC ATATGGAGCA  
1560  
TTTTTCAGAG TTTCTAACC CAGTCTTATT TTTCTAGTCA GTAAACATTI GTAAAAATAC  
1620  
50 TGTTTCACTA ATACTTACTG TTAACGTCTT TGAGAGAAAA GAAAAATATG AGAGAACTAT  
1680  
TGTTTGGGGA AGTTCAAGTG ATCTTTCAAT ATCATTACTA ACTTCTTCCA CTTTTTCCAA  
1740  
AATTTGAATA TTAACGCTAA AGGTGTAAGA CTTCAGATTI CAAATTAATC TTTCTATATT  
1800  
55 TTTTAAATTT ACAGAATATT ATATAACCCA CTGCTGAAAA AGAAAAAAT GATTGTTTTA  
1860  
GAAGTTAAAG TCAATATTGA TTTTAAATAT AAGTAATGAA GGCATATTTT CAATAACTAG  
1920  
TGATATGGCA TCGTTGCATT TTACAGTATC TTCAAAAATA CAGAATTTAT AGAATAATTT  
1980  
60 CTCCTCATTT AATATTTTTT AAAATCAAAG TTATGGTTTC CTCATTTTAC TAAAATCGTA  
2040  
TTCTAATTCT TCATTATAGT AAATCTATGA GCAACTCCTT ACTTCGGTTC CTCTGATTTT  
2100  
65 AAGGCCATAT TTTAAAAAAT CAAAAGGCAC TGTGAACAT TTTGAAGAAA ACACAACATT  
2160  
TTAATACAGA TTGAAAGGAC CTCTTCTGAA GCTAGAAACA ATCTATAGTT ATACATCTTC  
2220

ATTAATACTG TGTACCTTT TAAATAGTA ATTTTITACA TTTTCCTGTG TAAACCTAAT  
 2280  
 TGTGGTAGAA ATTTTACCA ACTCTATACT CAATCAAGCA AAATTCTGT ATATTCCCTG  
 2340  
 5 TGGAAATGTAC CTATGTGAGT TTCAGAAAT CTCAAAATAC GTGTTCAAAA ATTTCTGCTT  
 2400  
 TTGCATCTTT GGGACACCTC AGAAAACTTA TTAACAACCTG TGAATATGAG AAATACAGAA  
 2460  
 10 GAAAATAATA AGCCCTCTAT ACATAAATGC CCAGCACAAT TCATTGTTAA AAAACAACCA  
 2520  
 AACCTCACAC TACTGTATTT CATTATCTGT ACTGAAAGCA AATGCTTTGT GACTATTAA  
 2580  
 TGTTCACAT CATTCATTCA AAAAAAAAAA AAAAAAAAA  
 2618  
 15 2618  
  
 <210> 93  
 20 <211> 302  
 <212> ADN  
 <213> Homo sapiens  
  
 <400> 93  
 25 tgtgtcaacc tgaacaagct agaaccata gcaacagaag tctggctcat caacaagtc 60  
 atggagctgc tggagtagag tggcgtgacc tcagctcact gcaacctctg cctcctgagt  
 120  
 tcaagtgatt ctcgtgcctc agcctcccaa gtagctggga ttacagctcc tgccaccacg  
 180  
 30 cccggggctg gtatttgtgt cactggaatt actccaggca gcattgagct gccccatcat  
 240  
 gtcaagtaca agatccgaat ggacattgac aatgtggaga ggacaaataa aatcaaggat  
 300  
 gg  
 35 302  
  
 <210> 94  
 <211> 9593  
 40 <212> ADN  
 <213> Homo sapiens  
  
 <400> 94  
 caaacatgtc agctgttact ggaagtggcc tggcctctat ttatcttctt gatcctgac 60  
 45 tctgttcggc tgagctaccc accctatgaa caacatgaat gccatcttcc aaataaagcc  
 120  
 atgccctctg caggaacact tccttgggtt caggggatta tctgtaatgc caacaacccc  
 180  
 50 tgtttccggt acccgactcc tggggaggct cccggagttg ttggaaactt taacaaatcc  
 240  
 attgtggctc gcctgttctc agatgctcgg aggcttctt tatacagcca gaaagacacc  
 300  
 agcatgaagg acatgcgcaa agttctgaga acattacagc agatcaagaa atccagctca  
 360  
 55 aacttgaagc ttcaagattt cctggtggac aatgaaacct tctctgggtt cctgtatcac  
 420  
 aacctctctc tcccaaagtc tactgtggac aagatgctga gggctgatgt cattctccac  
 480  
 aaggatattt tgcaaggcta ccagttacat ttgacaagtc tgtgcaatgg atcaaaatca  
 540  
 60 gaagagatga ttcaacttgg tgaccaagaa gtttctgagc tttgtggcct accaagggag  
 600  
 aaactggctg cagcagagcg agtacttctt tccaacatgg acatcctgaa gccaatcctg  
 660  
 65 agaacactaa actctacatc tcccttcccg agcaaggagc tggccgaagc cacaaaaaca  
 720  
 ttgctgcata gtcttgggac tctggcccag gagctgttca gcatgagaag ctggagtgac  
 780

atgcgacagg aggtgatgtt tctgaccaat gtgaacagct ccagctcctc cacccaaatc  
 840  
 taccaggctg tgtctcgtat tgtctgcggg catcccaggg gaggggggct gaagatcaag  
 900  
 5 tctctcaact ggtatgagga caacaactac aaagccctct ttggaggcaa tggcactgag  
 960  
 gaagatgctg aaaccttcta tgacaactct acaactcctt actgcaatga tttgatgaag  
 1020  
 10 aatttggagt ctagtcctct tccccgatt atctggaaag ctctgaagcc gctgctcgtt  
 1080  
 gggaagatcc tgtatacacc tgacactcca gccacaaggc aggtcatggc tgaggtgaac  
 1140  
 aagaccttcc aggaactggc tgtgttccat gatctggaag gcatgtggga ggaactcagc  
 1200  
 15 cccaagatct ggaccttcat ggagaacagc caagaaatgg acctgttccg gatgctgttg  
 1260  
 gacagcaggg acaatgacca cttttgggaa cagcagttgg atggcttaga ttggacagcc  
 1320  
 20 caagacatcg tggcgTTTTT ggccaagcac ccagaggatg tccagtcag taatggttct  
 1380  
 gtgtacacct ggagagaagc tttcaacgag actaaccagg caatccggac catatctcgc  
 1440  
 ttcattggagt gtgtcaacct gaacaagcta gaaccatag caacagaagt ctggctcctc  
 1500  
 25 aacaagtcca tggagctgct ggagtacagt ggcgtgacct cagctcactg caacctctgc  
 1560  
 ctcttgagtt caagtattc tctgtcctca gcctcccaag tagctgggat tacagctcct  
 1620  
 gccaccacgc ccggggctgg tattgtgttc actggaatta ctccaggcag cattgagctg  
 1680  
 30 ccccatcatg tcaagtacaa gatccgaatg gacattgaca atgtggagag gacaaataaa  
 1740  
 atcaaggatg ggtactggga ccttggtcct cgagctgacc cctttgagga catgcggtac  
 1800  
 35 gtctgggggg gcttcgccta cttgcaggat gtgggtggagc aggcaatcat cagggtgctg  
 1860  
 acgggcaccg agaagaaaac tgggtgtctat atgcaacaga tgccctatcc ctgttacgtt  
 1920  
 gatgacatct ttctgcgggt gatgagccgg tcaatgcccc tcttcatgac gctggcctgg  
 1980  
 40 atttactcag tggctgtgat catcaagggc atcgtgtatg agaaggaggc acggctgaaa  
 2040  
 gagaccatgc ggatcatggg cctggacaac agcatcctct ggtttagctg gttcattagt  
 2100  
 45 agcctcattc ctcttcttgt gagcgtggc ctgctagtgg tcatcctgaa gttaggaaac  
 2160  
 ctgctgccct acagtgatcc cagcgtgggt tttgtcttcc tgtccgtgtt tgctgtgggtg  
 2220  
 acaatcctgc agtgcttct gattagcaca ctcttctcca gagccaaact ggcagcagcc  
 2280  
 50 tgtgggggca tcatctactt cacgctgtac ctgccctacg tcctgtgtgt ggcattggcag  
 2340  
 gactacgtgg gcttcacact caagatcttc gctagcctgc tgtctcctgt ggcttttggg  
 2400  
 55 tttggctgtg agtactttgc cctttttgag gagcagggca ttggagtgca gtgggacaac  
 2460  
 ctgtttgaga gtccgtgga ggaagatggc ttcaatctca ccacttcggt ctccatgatg  
 2520  
 ctgtttgaca ccttctctta tgggggtgat acctgttaca ttgaggctgt ctttccaggc  
 2580  
 60 cagtacggaa ttcccaggcc ctggtatttt ccttgcacca agtcctactg gtttggcgag  
 2640  
 gaaagtgatg agaagagcca ccctggttcc aaccagaaga gaatatcaga aatctgcatg  
 2700  
 65 gaggaggaaac ccaccttctt gaagctgggc gtgtccattc agaacctggt aaaagtctac  
 2760  
 cgagatggga tgaaggtggc tgtcgtggc ctggcactga atttttatga gggccagatc  
 2820

acctccttcc tgggccacaa tggagcgggg aagacgacca ccatgtcaat cctgaccggg  
 2880  
 ttgttcccc cgacctcggg caccgcctac atcctgggaa aagacattcg ctctgagatg  
 2940  
 5 agcaccatcc ggcagaacct gggggctctgt ccccgagcata acgtgctgtt tgacatgctg  
 3000  
 actgtcgaag aacacatctg gttctatgcc cgcttgaaa ggcctctctga gaagcacgtg  
 3060  
 10 aaggcggaga tggagcagat ggccctggat gttggtttgc catcaagcaa gctgaaaagc  
 3120  
 aaaacaagcc agctgtcagg tggaatgcag agaaagctat ctgtggcctt ggcctttgtc  
 3180  
 gggggatcta aggttgtcat tctggatgaa cccacagctg gtgtggaccc ttactcccg  
 3240  
 15 aggggaatat gggagctgct gctgaaatac cgacaaggcc gcaccattat tctctctaca  
 3300  
 caccacatgg atgaagcgga cgtcctgggg gacaggattg ccatcatctc ccatgggaag  
 3360  
 ctgtgctgtg tgggctcctc cctgtttctg aagaaccagc tgggaacagg ctactacctg  
 3420  
 20 accttgggtca agaaagatgt ggaatcctcc ctacgttccct gcagaaacag tagtagcact  
 3480  
 gtgtcatacc tgaaaaagga ggacagtgtt tctcagagca gttctgatgc tggcctgggc  
 3540  
 25 agcgaccatg agagtgcac gctgaccatc gatgtctctg ctatctccaa cctcatcagg  
 3600  
 aagcatgtgt ctgaagcccg gctgggtgaa gacatagggc atgagctgac ctatgtgctg  
 3660  
 30 ccatatgaag ctgctaagga gggagccttt gtggaactct ttcattgagat tgatgaccgg  
 3720  
 ctctcagacc tgggcatttc tagttatggc atctcagaga cgacctgga agaaatattc  
 3780  
 ctcaagggtg ccgaagagag tgggggtggat gctgagacct cagatggtac cttgccagca  
 3840  
 35 agacgaaaca ggcgggcctt cggggacaag cagagctgtc ttcgcccgtt cactgaagat  
 3900  
 gatgtgctg atccaaatga ttctgacata gaccagaaat ccagagagac agacttgctc  
 3960  
 40 agtgggatgg atggcaaagg gtcctaccag gtgaaaggct ggaaacttac acagcaacag  
 4020  
 tttgtggccc ttttgtggaa gagactgcta attgccagac ggagtccgaa aggatttttt  
 4080  
 gctcagattg tcttgccagc tgtgtttgtc tgcattgccc ttgtgttcag cctgatcgtg  
 4140  
 45 ccaccctttg gcaagtaccc cagcctggaa ctccagccct ggatgtacaa cgaacagtac  
 4200  
 acatttgtca gcaatgatgc tcctgaggac acgggaaccc tggaaactctt aaacgccctc  
 4260  
 50 accaaagacc ctggcttcgg gacccgctgt atggaaggaa acccaatccc agacacgccc  
 4320  
 tgccaggcag gggaggaaga gtggaccact gcccagttc cccagaccat catggacctc  
 4380  
 ttccagaatg ggaactggac aatgcagaac ccttcacctg catgccagtg tagcagcgac  
 4440  
 55 aaaatcaaga agatgtgcc tgtgtgtccc ccaggggcag gggggctgcc tcctccacaa  
 4500  
 agaaaacaaa acactgcaga tctccttcag gacctgacag gaagaaacat ttcggattat  
 4560  
 ctggtgaaga cgtatgtgca gatcatagcc aaaagcttaa agaacaagat ctgggtgaat  
 4620  
 60 gagttaggt atggcggtt tccctgggt gtcagtaata ctcaagcact tcctccgagt  
 4680  
 caagaagtta atgatgccac caaacaatg aagaaacacc taaagctggc caaggacagt  
 4740  
 65 tctgcagatc gatttctcaa cagctggga agatttatga caggactgga caccagaaat  
 4800  
 aatgtcaagg tgtggtcaa taacaagggc tggcatgcaa tcagctctt cctgaatgtc  
 4860



atcaacaatg ccattctccg ggccaacctg caaaaggag agaaccctag ccattatgga  
4920  
attactgctt tcaatcatcc cctgaatctc accaagcagc agctctcaga ggtggctccg  
4980  
5 atgaccacat cagtggatgt ccttgtgtcc atctgtgtca tctttgcaat gtccttcgtc  
5040  
ccagccagct ttgtcgtatt cctgatccag gagcgggtca gcaaagcaaa acacctgcag  
5100  
10 ttcatcagtg gagtgaagcc tgtcatctac tggctctcta attttgtctg ggatatgtgc  
5160  
aattacgttg tccctgccac actggtcatt atcatcttca tctgcttcca gcagaagtc  
5220  
tatgtgtcct ccaccaatct gcctgtgcta gcccttctac ttttgcgtga tgggtgggtca  
5280  
15 atcacacctc tcatgtaccc agcctccttt gtgttcaaga tccccagcac agcctatgtg  
5340  
gtgctcacca gcgtgaacct cttcattggc attaatggca gcgtggccac ctttgtgctg  
5400  
gagctgttca ccgacaataa gctgaataat atcaatgata tcctgaagtc cgtgttcttg  
20 5460  
atcttccac atttttgcct gggacgaggg ctcatcgaca tggtgaaaaa ccaggcaatg  
5520  
gctgatgcc tggaaagggt tggggagaat cgctttgtgt caccattatc ttgggacttg  
5580  
25 gtgggacgaa acctcttcgc catggccgtg gaagggttg tgttcttcct cactactgtt  
5640  
ctgatccagt acagattctt catcaggccc agacctgtaa atgcaaagct atctcctctg  
5700  
aatgatgaag atgaagatgt gaggcgggaa agacagagaa ttcttgatgg tggaggccag  
30 5760  
aatgacatct tagaaatcaa ggagttgacg aagatatata gaaggaagcg gaagcctgct  
5820  
gttgacagga tttgcgtggg cttcctcct ggtgagtgt tgggctcct gggagttaat  
5880  
35 ggggctggaa aatcatcaac tttcaagatg ttaacaggag ataccactgt taccagagga  
5940  
gatgctttcc ttaacagaaa tagtatctta tcaaacatcc atgaagtaca tcagaacatg  
6000  
ggctactgcc ctcatgttga tgccatcaca gagctgttga ctgggagaga acacgtggag  
40 6060  
ttctttgccc ttttgagagg agtcccagag aaagaagttg gcaaggttgg tgagtgggcg  
6120  
attcggaaac tgggcctcgt gaagtatgga gaaaaatatg ctggtaacta tagtggaggc  
6180  
45 aacaaacgca agctctctac agccatggct ttgatcggcg ggcctcctgt ggtgtttctg  
6240  
gatgaacca ccacaggcat ggatcccaa gcccgcggt tcttgtggaa ttgtgcccta  
6300  
agtgttgtca aggaggggag atcagtagtg cttacatctc atagtatgga agaattgtgaa  
50 6360  
gctctttgca ctaggatggc aatcatggc aatggaaggt tcagggtcct tggcagtgtc  
6420  
cagcatctaa aaaatagggt tggagatggt tatacaatag ttgtacgaat agcagggtcc  
6480  
55 aaccgggacc tgaagcctgt ccaggatttc tttggacttg catttcttgg aagtgttcca  
6540  
aaagagaaac accggaacat gctacaatac cagcttccat cttcattatc ttctctggcc  
6600  
aggatattca gcatcctctc ccagagcaaa aagcgactcc acatagaaga ctactctgtt  
60 6660  
tctcagacaa cacttgacca agtatgttg aactttgcc aggaccaaag tgatgatgac  
6720  
cacttaaaag acctctcatt acacaaaaac cagacagtag tggacgttgc agttctcaca  
6780  
65 tcttttctac aggatgagaa agtgaaagaa agctatgat gaagaatcct gttcatacgg  
6840  
ggtggctgaa agtaaagagg nactagactt tcctttgcac catgtgaagt gttgtggaga  
6900

aaagagccag aagttgatgt gggaagaagt aaactggata ctgtactgat actattcaat  
 6960  
 gcaatgcaat tcaatgcaat gaaaacaaaa ttccattaca ggggcagtgc cttttagacc  
 7020  
 5 tatgtcttgt atggctctca agtgaaagac ttgaatttag ttttttacct atacctatgt  
 7080  
 gaaactctat tatggaaccc aatggacata tgggtttgaa ctcacacttt tttttttttt  
 7140  
 10 ttgttctctgt gtattctcat tggggttgca acaataattc atcaagtaat catggccagc  
 7200  
 gattattgat caaaatcaaa aggtaatgca catctcatt cactaagcca tgccatgccc  
 7260  
 aggagactgg tttcccggtg acacatccat tgctggcaat gagtgtgcca gagttattag  
 7320  
 15 tgccaagttt ttcagaaagt ttgaagcacc atggtgtgtc atgtcactt ttgtgaaagc  
 7380  
 tgctctgctc agagtctatc aacattgaat atcagttgac agaatggtgc catgctgtgc  
 7440  
 20 taacatcctg ctttgattcc ctctgataag ctgttctggt ggcagtaaca tgcaacaaaa  
 7500  
 atgtgggtgt ctctaggcac gggaaacttg gttccattgt tatattgtcc tatgcttcga  
 7560  
 gccatgggtc tacagggtca tccttatgag actcttaaat atacttagat cctggtgaaga  
 7620  
 25 ggcaagaat caacagccaa actgctgggg ctgcaagctg ctgaagccag ggcattggat  
 7680  
 taaagagatt gtgcgttcaa acctaggga gctgtgccc atttgtcctg actgtctgct  
 7740  
 aacatggtac actgcctc cagatgttta tctgacacaa gtgtattatt tctggctttt  
 7800  
 30 tgaaattaatc tagaaaatga aaagatggag ttgtattttg aaaaaaatgt ttgtactttt  
 7860  
 taatgttatt tggaatttta agttctatca gtgacttctg aatccttaga atggcctctt  
 7920  
 35 tgtagaaccc tgtggtatag aggagtatgg ccaactgccc actattttta ttttcttatg  
 7980  
 taagtttgca tatcagtcac gactagtgc tagaaagcaa tgtgatggtc aggatctcat  
 8040  
 40 gacattatat ttgagtttct ttcagatcat ttaggatact cttaatctca cttcatcaat  
 8100  
 caaatatttt ttgagtgtat gctgtagctg aaagagtatg tacgtacgta taagactaga  
 8160  
 gagatattaa gtctcagtac acttctctgt ccatgttatt cagctcactg gtttacaat  
 8220  
 45 ataggttgtc ttgtggttgt aggagccac tgtaacaata ctgggcagcc tttttttttt  
 8280  
 tttttaattg caacaatgca aaagccaaga aagtataagg gtcacaagtc taaacaatga  
 8340  
 attcttcaac agggaaaaca gctagcttga aaacttgctg aaaaacacaa cttgtgttta  
 8400  
 50 tggcatttag taccttcaaa taattggctt tgcagatatt ggatacccca ttaaatctga  
 8460  
 cagtctcaaa ttttctatct cttcaatcac tagtcaagaa aatatataaa acaacaaata  
 8520  
 55 cttccatag gagcattttt cagagttttc taaccagtc ttatttttct agtcagtaaa  
 8580  
 catttgtaaa aatactgttt cactaatact tactgttaac tgtcttgaga gaaaagaaaa  
 8640  
 60 atatgagaga actattgttt ggggaagttc aagtgtctt tcaatatcat tactaacttc  
 8700  
 ttccactttt tccaaaattt gaatattaac gctaaagggt taagacttca gatttcaaat  
 8760  
 taatctttct atatttttta aatttacaga atattatata acccactgct gaaaaagaaa  
 8820  
 65 aaaaatgattg ttttagaagt taaagtcaat attgatttta aatataagta atgaaggcat  
 8880  
 atttccaata actagtata tggcatcggt gcattttaca gtatcttcaa aaatacagaa  
 8940

tttatagaat aatttctcct catttaatat ttttcaaat caaagttatg gtttctcat  
 9000  
 tttactaaaa tcgtattcta attcttcatt atagtaaate tatgagcaac tccttacttc  
 9060  
 5 ggttcctctg atttcaaggc catattttta aaaatcaaaa ggcactgtga actattttga  
 9120  
 agaaaaacaca acattttaat acagattgaa aggacctctt ctgaagctag aaacaatcta  
 9180  
 tagttataca tcttcattaa tactgtgtta ctttttaaaa tagtaatttt ttacattttc  
 9240  
 10 ctgtgtaaac ctaattgtgg tagaaatttt taccaactct atactcaate aagcaaaatt  
 9300  
 tctgtatatt cctgtggaa tgtacctatg tgagtctcag aaattctcaa aatacgtgtt  
 9360  
 15 caaaaatttc tgcttttgca tctttgggac acctcagaaa acttattaac aactgtgaat  
 9420  
 atgagaaata cagaagaaaa taataagccc tctatacata aatgccagc acaattcatt  
 9480  
 gttaaaaaac aaccaaact cacactactg tatttcatta tctgtactga aagcaaatgc  
 9540  
 20 tttgtgacta ttaaatgttg cacatcattc attcaaaaaa aaaaaaaaaa aaa  
 9593

25 <210> 95  
 <211> 173  
 <212> ADN  
 <213> Homo sapiens

30 <400> 95  
 ctgggacctt ggtcctcgag ctgacctctt tgaggacatg cggtagctct gggggggctt 60  
 cgcctacttg caggatgtgg tggagcaggc aatcatcagg gtgctacggg caccgagaag  
 120  
 aaaactgggtg tctatatgca acagatgcc taccctgtt acgttgatga cat  
 35 173

40 <210> 96  
 <211> 9495  
 <212> ADN  
 <213> Homo sapiens

45 <400> 96  
 caaacatgtc agctgttact ggaagtggcc tggcctctat ttatcttctt gatcctgac 60  
 tctgttcggc tgagctaccc accctatgaa caacatgaat gccattttcc aaataaagcc  
 120  
 atgccctctg caggaacact tccttgggtt caggggatta tctgtaatgc caacaacccc  
 180  
 tgtttccggt acccgactcc tggggaggct cccggagttg ttggaaactt taacaaatcc  
 240  
 50 attgtggctc gcctgttctc agatgctcgg aggtctcttt tatacagcca gaaagacacc  
 300  
 agcatgaagg acatgcgcaa agttctgaga acattacagc agatcaagaa atccagctca  
 360  
 55 aacttgaagc ttcaagattt cctgggtggac aatgaaacct tctctgggtt cctgtatcac  
 420  
 aacctctctc tcccaaagtc tactgtggac aagatgctga gggctgatgt cattctccac  
 480  
 aaggtatttt tgcaaggcta ccagttacat ttgacaagtc tgtgcaatgg atcaaaatca  
 540  
 60 gaagagatga ttcaacttgg tgaccaagaa gtttctgagc tttgtggcct accaaggagg  
 600  
 aaactggctg cagcagagcg agtacttcgt tccaacatgg acatcctgaa gccaatcctg  
 660  
 65 agaacactaa actctacatc tcccttcccg agcaaggagc tggccgaagc cacaaaaaca  
 720  
 ttgctgcata gtcttgggac tctggcccag gagctgttca gcatgagaag ctggagtgac  
 780

atgcgacagg aggtgatggt tctgaccaat gtgaacagct ccagctcctc cacccaaatc  
840  
taccaggctg tgtctcgat tgtctcgagg catcccgagg gaggggggct gaagatcaag  
900  
5 tctctcaact ggtatgagga caacaactac aaagccctct ttggaggcaa tggcactgag  
960  
gaagatgctg aaaccttcta tgacaactct acaactcctt actgcaatga tttgatgaag  
1020  
10 aatttggagt ctagtcctct tccccgatt atctggaaag ctctgaagcc gctgctcgtt  
1080  
gggaagatcc tgtatacacc tgacactcca gccacaaggc aggtcatggc tgaggatgaac  
1140  
aagaccttcc aggaactggc tgtgttccat gatctggaag gcatgtggga ggaactcagc  
1200  
15 cccaagatct ggaccttcat ggagaacagc caagaaatgg accttgtccg gatgctgttg  
1260  
gacagcaggg acaatgacca cttttgggaa cagcagttgg atggcttaga ttggacagcc  
1320  
20 caagacatcg tggcgttttt ggccaagcac ccagaggatg tccagtccag taatggttct  
1380  
gtgtacacct ggagagaagc tttcaacgag actaaccagg caatccggac catatctcgc  
1440  
ttcatggagt gtgtcaacct gaacaagcta gaacccatag caacagaagt ctggctcatc  
1500  
25 aacaagtcca tggagctgct ggatgagagg aagttctggg ctggtattgt gttcactgga  
1560  
attactccag gcagcattga gctgccccat catgtcaagt acaagatccg aatggacatt  
1620  
gacaatgtgg agaggacaaa taaaatcaag gatgggtact gggaccctgg tcctcgagct  
30 1680  
gaccctttg aggacatgag gtacgtctgg gggggcttcg cctacttgca ggatgtggtg  
1740  
gagcaggcaa tcatcagggt gctcgggcac cgagaagaaa actggtgtct atatgcaaca  
1800  
35 gatgccctat ccctgttacg ttgatgacat ctttctcgcg gtgatgagcc ggtcaatgcc  
1860  
cctcttcatg acgtggcct ggatttactc agtggctgtg atcatcaagg gcatcgtgta  
1920  
tgagaaggag gcacggctga aagagaccat gcgatcatg ggcctggaca acagcatcct  
40 1980  
ctggtttagc tggttcatta gtagcctcat tcctcttctt gtgagcgtg gcctgctagt  
2040  
ggtcacctg aagttaggaa acctgctgcc ctacagtgt cccagcgtgg tgtttgtctt  
2100  
45 cctgtccgtg tttgctgtgg tgacaactct gcagtgttc ctgattagca cactcttctc  
2160  
cagagccaac ctggcagcag cctgtgggg catcatctac ttcacgtgt acctgcccta  
2220  
50 cgtcctgtgt gtggcatggc aggactacgt gggcttcaca ctcaagatct tcgctagcct  
2280  
gctgtctcct gtggcttttg ggtttggctg tgagtacttt gccctttttg aggagcaggg  
2340  
cattggagtg cagtgggaca acctgtttga gagtccctgt gaggaagatg gcttcaatct  
2400  
55 caccacttcg gtctccatga tgctgtttga caccctctc tatggggtga tgacctgga  
2460  
cattgaggct gtctttccag gccagtacgg aattcccagg ccttgggtatt ttccttgac  
2520  
caagtcctac tggtttgccg aggaagtgga tgagaagagc caccctggtt ccaaccagaa  
60 2580  
gagaatatca gaaatctgca tggaggagga acccaccac ttgaagctgg gcgtgtccat  
2640  
tcagaacctg gtaaaagtct accgagatgg gatgaaggtg gctgtcgatg gcctggcact  
2700  
65 gaatttttat gagggccaga tcacctcctt cctgggccac aatggagcgg ggaagacgac  
2760  
caccatgtca atcctgaccg ggtgttccc cccgacctcg ggcaccgct acatcctggg  
2820

aaaagacatt cgctctgaga tgagcaccat ccggcagaac ctgggggtct gtccccagca  
2880  
taacgtgctg tttgacatgc tgactgtcga agaacacatc tggttctatg cccgcttgaa  
2940  
5 agggctctct gagaagcacg tgaaggcgga gatggagcag atggccctgg atgttggttt  
3000  
gccatcaagc aagctgaaaa gcaaaacaag ccagctgtca ggtggaatgc agagaaagct  
3060  
10 atctgtggcc ttggcctttg tcgggggata taaggttgte attctggatg aaccacagc  
3120  
tggtgtggac cttactccc gcagggaat atgggagctg ctgctgaaat accgacaagg  
3180  
ccgcaccatt attctctcta cacaccacat ggatgaagcg gacgtcctgg gggacaggat  
3240  
15 tgccatcatc tcccatggga agctgtgctg tgtgggctcc tccctgttcc tgaagaacca  
3300  
gctgggaaca ggctactacc tgaccttggc caagaaagat gtggaatcct ccctcagttc  
3360  
20 ctgcagaaac agtagtagca ctgtgtcata cctgaaaaag gaggacagtg tttctcagag  
3420  
cagttctgat gctggcctgg gcagcgacca tgagagtac acgctgacca tcgatgtctc  
3480  
tgctatctcc aacctcatca ggaagcatgt gtctgaagcc cggctggtgg aagacatagg  
3540  
25 gcatgagctg acctatgtgc tgccatatga agctgctaag gagggagcct ttgtggaact  
3600  
ctttcatgag attgatgacc ggctctcaga cctgggcatt tctagttagt gcatctcaga  
3660  
30 gacgacctg gaagaaatat tctcaaggt ggccgaagag agtggggtgg atgctgagac  
3720  
ctcagatggt accttgccag caagacgaaa caggcgggcc ttcggggaca agcagagctg  
3780  
tcttcgcccg ttcactgaag atgatgctgc tgatccaaat gattctgaca tagaccaga  
3840  
35 atccagagag acagacttgc tcagtgggat ggatggcaaa gggccctacc aggtgaaagg  
3900  
ctggaaactt acacagcaac agtttgtggc cttttgtgg aagagactgc taattgccag  
3960  
40 acggagtcgg aaaggatttt ttgctcagat tgtcttgcca gctgtgtttg tctgcattgc  
4020  
ccttgtgttc agcctgatcg tgccaccctt tggcaagtac cccagcctgg aacttcagcc  
4080  
ctggatgtac aacgaacagt acacatttgt cagcaatgat gctcctgagg acacgggaac  
4140  
45 cctggaactc ttaaagcccc tcaccaaaaga ccctggcttc gggaccgct gtatggaagg  
4200  
aaaccaatc ccagacacgc cctgccaggc aggggaggaa gattggacca ctgccccagt  
4260  
50 tccccagacc atcatggacc tcttcagaa tgggaactgg acaatgcaga acccttcacc  
4320  
tgcatgccag tgtagcagcg acaaaatcaa gaagatgctg cctgtgtgtc cccaggggc  
4380  
aggggggctg cctcctccac aaagaaaaca aaacactgca gatatccttc aggacctgac  
4440  
55 aggaagaaac atttcggatt atctggtgaa gacgtatgtg cagatcatag ccaaaagctt  
4500  
aaagaacaag atctgggtga atgagtttag gtatggcggc tttccctgg gtgtcagtaa  
4560  
60 tactcaagca cttcctccga gtcaagaagt taatgatgcc accaaacaaa tgaagaaaca  
4620  
cctaaagctg gccaaggaca gttctgcaga tcgatttctc aacagcttgg gaagatttat  
4680  
gacaggactg gacaccagaa ataatgtcaa ggtgtggttc aataacaagg gctggcatgc  
4740  
65 aatcagctct ttcctgaatg tcatcaacaa tgccattctc cgggccaacc tgcaaaaggg  
4800  
agagaaccct agccattatg gaattactgc tttcaatcat cccctgaatc tcaccaagca  
4860

gcagctctca gaggtggctc cgatgaccac atcagtggtat gtccttgtgt ccatctgtgt  
 4920  
 catctttgca atgtccttcg tcccagccag ctttgtcgta ttcctgatcc aggagcgggt  
 4980  
 5 cagcaaagca aaacacctgc agttcatcag tggagtgaag cctgtcatct actggctctc  
 5040  
 taattttgtc tgggatatgt gcaattacgt tgtccctgcc aacttggtca ttatcatctt  
 5100  
 catctgcttc cagcagaagt cctatgtgtc ctccaccaat ctgcctgtgc tagcccttct  
 10 5160  
 acttttgctg tatgggtggt caatcacacc tctcatgtac ccagcctcct ttgtgttcaa  
 5220  
 gatccccagc acagcctatg tgggtgctac cagcgtgaac ctcttcattg gcattaatgg  
 5280  
 15 cagcgtggcc acctttgtgc tggagctggt caccgacaat aagctgaata atatcaatga  
 5340  
 tatcctgaag tccgtgttct tgatcttccc acatttttgc ctgggacgag ggctcatcga  
 5400  
 catgggtgaaa aaccaggcaa tggctgatgc cctggaaaag tttggggaga atcgctttgt  
 20 5460  
 gtcaccatta tcttgggact tgggtgggacg aaacctcttc gccatggccg tggaggggt  
 5520  
 ggtgttcttc ctcatctactg ttctgatcca gtacagattc ttcacagggc ccagacctgt  
 5580  
 25 aaatgcaaag ctatctcttc tgaatgatga agatgaagat gtgaggcggg aaagacagag  
 5640  
 aattcttgat ggtggaggcc agaatgacat cttagaaatc aaggagtga cgaagatata  
 5700  
 tagaaggaa cgggaagcctg ctgttgacag gatttgcgtg ggcatctctc ctggtgagtg  
 30 5760  
 ctttgggctc ctgggagtta atggggctgg aaaatcatca actttcaaga tgtaaacagg  
 5820  
 agataccact gttaccagag gagatgcttt ccttaacaga aatagtatct tatcaaact  
 5880  
 35 ccatgaagta catcagaaca tgggctactg cctcagttt gatgccatca cagagctggt  
 5940  
 gactgggaga gaacacgtgg agttctttgc ccttttgaga ggagtcccag agaaagaagt  
 6000  
 tggcaaggtt ggtgagtggg cgattcggaa actgggcctc gtgaagtatg gagaaaaata  
 40 6060  
 tgctggtaac tatagtggag gcaacaaacg caagctctct acagccatgg ctttgatcgg  
 6120  
 cgggcctcct gtggtgttct tggatgaacc caccacaggc atggatccca aagcccgcg  
 6180  
 45 gttcttgtgg aattgtgccc taagtgttgt caaggagggg agatcagtag tgcttacatc  
 6240  
 tcatagtatg gaagaatgtg aagctctttg cactaggatg gcaatcatgg tcaatggaag  
 6300  
 gttcaggtgc cttggcagtg tccagcatct aaaaaatagg tttggagatg gttatacaat  
 50 6360  
 agttgtacga atagcagggt ccaacccgga cctgaagcct gtccaggatt tctttggact  
 6420  
 tgcatttctt ggaagtgttc caaaagagaa acaccggaac atgctacaat accagcttcc  
 6480  
 55 atctttcatia tcttctcttg ccaggatatt cagcatcttc tcccagagca aaaagcgact  
 6540  
 ccacatagaa gactactctg tttctcagac aacacttgac caagtatttg tgaactttgc  
 6600  
 caaggaccaa agtgatgatg accacttaaa agacctctca ttacacaaaa accagacagt  
 60 6660  
 agtggacgtt gcagttctca catcttttct acaggatgag aaagtgaag aaagctatgt  
 6720  
 atgaagaatc ctgttcatac ggggtggctg aaagtaaaga gmnactagac tttcctttgc  
 6780  
 65 accatgtgaa gtgttgtgga gaaaagagcc agaagttgat gtgggaagaa gtaaaactgga  
 6840  
 tactgtactg atactattca atgcaatgca attcaatgca atgaaaacaa aattccatta  
 6900

caggggcagt gcctttgtag cctatgtctt gtagggctct caagtgaag acttgaattt  
 6960  
 agttttttac ctataacctat gtgaaactct attatggaac ccaatggaca tatgggtttg  
 7020  
 5 aactcacact tttttttttt ttttgttctt gtgtattctc attgggggtg caacaataat  
 7080  
 tcatcaagta atcatggcca gcgattattg atcaaatca aaaggtaatg cacatcctca  
 7140  
 10 ttcactaagc catgccatgc ccaggagact gggttcccg tgacacatcc attgctggca  
 7200  
 atgagtgtgc cagagttatt agtgccaagt ttttcagaaa gtttgaagca ccatgggtgtg  
 7260  
 tcatgtctac ttttgtgaaa gctgctctgc tcagagtcta tcaacattga atatcagttg  
 7320  
 15 acagaatggt gccatgcgtg gctaacatcc tgctttgatt ccctctgata agctgttctg  
 7380  
 gtggcagtaa catgcaacaa aaatgtgggt gtctctaggc acgggaaact tggttccatt  
 7440  
 gttatattgt cctatgtctc gagccatggg tctacagggt catccttatg agactcttaa  
 7500  
 20 atatacttag atcctggtaa gaggcaaaga atcaacagcc aaactgctgg ggctgcaagc  
 7560  
 tgctgaagcc agggcatggg attaaagaga ttgtgcgttc aaacctaggg aagcctgtgc  
 7620  
 25 ccatttgtcc tgactgtctg ctaacatggt acactgcac tcaagatgtt tatctgacac  
 7680  
 aagtgtatta tttctggctt tttgaattaa tctagaaaat gaaaagatgg agttgtattt  
 7740  
 tgacaaaaat gtttgtactt tttaatgtta tttggaattt taagttctat cagtgaactc  
 7800  
 30 tgaatcctta gaatggcctc tttgtagaac cctgtggtat agaggagtat ggccactgcc  
 7860  
 ccaactatctt tattttctta tgtaagtttg catatcagtc atgactagtg cctagaaagc  
 7920  
 35 aatgtgatgg tcaggatctc atgacattat atttgagttt ctttcagatc atttaggata  
 7980  
 ctcttaatct cacttcatca atcaaatatt ttttgagtgt atgctgtagc tgaaagagta  
 8040  
 tgtacgtacg tataagacta gagagatatt aagtctcagt acacttcctg tgccatgtta  
 8100  
 40 ttcagctcac tgggtttaca atataggttg tcttgtggtt gtaggagccc actgtaacaa  
 8160  
 tactgggcag cctttttttt tttttttaat tgcaacaatg caaaagccaa gaaagtataa  
 8220  
 45 gggtcacaag tctaacaat gaattcttca acagggaaaa cagctagctt gaaaacttgc  
 8280  
 tgaaaaacac aacttgtgtt tatggcattt agtaccttca aataattggc tttgcagata  
 8340  
 ttggataccc cattaaatct gacagtctca aatttttcat ctcttcaatc actagtcaag  
 8400  
 50 aaaaaataaa aaacaacaaa tacttccata tggagcattt ttcagagttt tctaaccag  
 8460  
 tcttattttt ctagtacgta aacatttcta aaaatactgt ttcactaata cttactgtta  
 8520  
 55 actgtcttga gagaaaagaa aaatatgaga gaactattgt ttggggaagt tcaagtgatc  
 8580  
 tttcaatatc attactaact tcttccactt ttccaaaat ttgaatatta acgctaaagg  
 8640  
 tgtaagactt cagatttcaa attaatcttt ctatatTTTT taaatttaca gaatattata  
 8700  
 60 taaccactg ctgaaaaaga aaaaaatgat tgtttttagaa gttaaagtca atattgattt  
 8760  
 taaatataag taatgaaggc atatttccaa taactagtga tatggcatcg ttgcatttta  
 8820  
 65 cagtatcttc aaaaatacag aatttataga ataatttctc ctcaattaat atttttcaaa  
 8880  
 atcaaaagta tggtttctc attttactaa aatcgtattc taattcttca ttatagtaaa  
 8940

tctatgagca actccttact tcggttcctc tgatttcaag gccatatttt aaaaaatcaa  
 9000  
 aaggcactgt gaactatttt gaagaaaaca caacatttta atacagattg aaaggacctc  
 9060  
 5 tttctgaagct agaaacaatc tatagttata catcttcatt aatactgtgt taccttttaa  
 9120  
 aatagtaatt ttttacattt tcctgtgtaa acctaattgt ggtagaaatt ttaccaact  
 9180  
 ctatactcaa tcaagcaaaa tttctgtata ttccctgtgg aatgtaccta tgtgagtttc  
 10 9240  
 agaaattctc aaaatacgtg ttcaaaaatt tctgcttttg catctttggg acacctcaga  
 9300  
 aaacttatta acaactgtga atatgagaaa tacagaagaa aataataagc cctctatata  
 9360  
 15 taaatgccca gcacaattca ttgttaaaaa acaaccaaac ctacactac tgtatttcat  
 9420  
 tatctgtact gaaagcaaat gctttgtgac tattaatgt tgcacatcat tcattcaaaa  
 9480  
 aaaaaaaaaa aaaaa  
 20 9495

<210> 97  
 <211> 41  
 25 <212> ADN  
 <213> Homo sapiens

<400> 97  
 30 tgaaggctgt tcttctatca gtgtgtcaac ctgaacaagc t 41

<210> 98  
 <211> 41  
 <212> ADN  
 35 <213> Homo sapiens

<400> 98  
 tgaaggctgt tcttctatca atgtgtcaac ctgaacaagc t 41

40 <210> 99  
 <211> 41  
 <212> ADN  
 <213> Homo sapiens

45 <400> 99  
 agttacctgc aagccactgt ttttaaccag tttatactgt g 41

50 <210> 100  
 <211> 41  
 <212> ADN  
 <213> Homo sapiens

55 <400> 100  
 agttacctgc aagccactgt atttaaccag tttatactgt g 41

60 <210> 101  
 <211> 41  
 <212> ADN  
 <213> Homo sapiens

65 <400> 101  
 caggctcaga ggccttgcc catcacctg gctcacgtgt g 41

<210> 102



<211> 41  
<212> ADN  
<213> Homo sapiens

5 <400> 102  
caggctcaga ggccttgcc tatcacctg gctcacgtg g 41

10 <210> 103  
<211> 41  
<212> ADN  
<213> Homo sapiens

15 <400> 103  
atgggcctgg acaacagcat cctctggtt agctggttca t 41

20 <210> 104  
<211> 41  
<212> ADN  
<213> Homo sapiens

25 <400> 104  
atgggcctgg acaacagcat actctggtt agctggttca t 41

30 <210> 105  
<211> 41  
<212> ADN  
<213> Homo sapiens

35 <400> 105  
aaggaggag aagaagaaa aaaatccaag cctctggtag a 41

40 <210> 106  
<211> 41  
<212> ADN  
<213> Homo sapiens

45 <400> 106  
aaggaggag aagaagaaa gaaatccaag cctctggtag a 41

50 <210> 107  
<211> 41  
<212> ADN  
<213> Homo sapiens

55 <400> 107  
tcttccttt gcagagacac gcctgccag gcaggggagg a 41

60 <210> 108  
<211> 41  
<212> ADN  
<213> Homo sapiens

65 <400> 108  
tcttccttt gcagagacac accctgccag gcaggggagg a 41

<210> 109  
<211> 23  
<212> ADN  
<213> Homo sapiens

<400> 109

	ccttgctcc tagttagga ttt	23
5	<210> 110 <211> 24 <212> ADN <213> Homo sapiens	
10	<400> 110 aatataatag gtgctctgga cctc	24
15	<210> 111 <211> 25 <212> ADN <213> Homo sapiens	
20	<400> 111 atacaaaaat agaaaaaggg gcttg	25
25	<210> 112 <211> 25 <212> ADN <213> Homo sapiens	
30	<400> 112 atggatgaga aggaaagagg ttac	25
35	<210> 113 <211> 25 <212> ADN <213> Homo sapiens	
40	<400> 113 cattggatca tacgtacatt tcaga	25
45	<210> 114 <211> 25 <212> ADN <213> Homo sapiens	
50	<400> 114 tcactttccc caactataaa tggat	25
55	<210> 115 <211> 25 <212> ADN <213> Homo sapiens	
60	<400> 115 gtagatcata caagtgagtg cttgg	25
65	<210> 116 <211> 25 <212> ADN <213> Homo sapiens	
	<400> 116 ctgttctcaa cttgctgctt ttatt	25
	<210> 117 <211> 23	

<212> ADN  
<213> Homo sapiens

<400> 117  
5 gcaaattcaa atttctccag gta 23

<210> 118  
<211> 23  
10 <212> ADN  
<213> Homo sapiens

<400> 118  
15 gcacaaagaa aggacatcag cta 23

<210> 119  
<211> 23  
20 <212> ADN  
<213> Homo sapiens

<400> 119  
cagtgcttac ccctgcta atc 23

25

<210> 120  
<211> 23  
<212> ADN  
<213> Homo sapiens

30

<400> 120  
gagatggaga aatcattcac agc 23

35

<210> 121  
<211> 23  
<212> ADN  
<213> Homo sapiens

40

<400> 121  
acatgtggaa tgacctaaac acc 23

45

<210> 122  
<211> 23  
<212> ADN  
<213> Homo sapiens

50

<400> 122  
cttaggacat ttggccttgc tat 23

55

<210> 123  
<211> 24  
<212> ADN  
<213> Homo sapiens

<400> 123  
60 catttctgtt ttaagagcct gtca 24

65

<210> 124  
<211> 23  
<212> ADN  
<213> Homo sapiens

<400> 124  
AATGTGGCAT GCAGTTGATA AAT\_ 24

<210> 125  
<211> 23  
5 <212> ADN  
<213> Homo sapiens  
  
<400> 125  
10 gtttgtggtt gttacggaat gat 23  
  
<210> 126  
<211> 23  
<212> ADN  
15 <213> Homo sapiens  
  
<400> 126  
CCTCCCAACA TGATATCTCA CTC\_ 24  
20  
  
<210> 127  
<211> 23  
<212> ADN  
25 <213> Homo sapiens  
  
<400> 127  
gtctgggacc tgtagtcagg ttt 23  
  
<210> 128  
<211> 23  
<212> ADN  
30 <213> Homo sapiens  
  
<400> 128  
35 ccaatagaca gaatcaggcc ata 23  
  
<210> 129  
<211> 23  
<212> ADN  
40 <213> Homo sapiens  
  
<400> 129  
45 tgccaacatt tattagagga agc 23  
  
<210> 130  
<211> 23  
50 <212> ADN  
<213> Homo sapiens  
  
<400> 130  
atccgtttaa cctgccaact act 23  
55  
  
<210> 131  
<211> 23  
<212> ADN  
60 <213> Homo sapiens  
  
<400> 131  
tctcaggagc cgtttattta atg 23  
65  
  
<210> 132  
<211> 23  
<212> ADN

<213> Homo sapiens

<400> 132  
5 gccaacttta ccatgagttg aaa 23

<210> 133  
<211> 23  
<212> ADN  
10 <213> Homo sapiens

<400> 133  
tctgatcata gtgttttgcc ttg 23

15 <210> 134  
<211> 23  
<212> ADN  
<213> Homo sapiens

20 <400> 134  
TGTTCCCTA CAATGAGATT CAC\_ 24

25 <210> 135  
<211> 22  
<212> ADN  
<213> Homo sapiens

30 <400> 135  
gggtgaacag atgtttttcc tt 22

<210> 136  
35 <211> 23  
<212> ADN  
<213> Homo sapiens

<400> 136  
40 tagctggaac atttcctgat gat 23

<210> 137  
<211> 23  
45 <212> ADN  
<213> Homo sapiens

<400> 137  
50 ccctttcttg tctgataatg gtg 23

<210> 138  
<211> 23  
<212> ADN  
55 <213> Homo sapiens

<400> 138  
CACAAATAAA CACTGTCCTC TGG\_ 24

60 <210> 139  
<211> 2201  
<212> PRT  
<213> Homo sapiens

65 <400> 139  
Met Pro Ser Ala Gly Thr Leu Pro Trp Val Gln Gly Ile Ile Cys Asn  
1 5 10 15

Ala Asn Asn Pro Cys Phe Arg Tyr Pro Thr Pro Gly Glu Ala Pro Gly  
                   20                  25                  30  
 5 Val Val Gly Asn Phe Asn Lys Ser Ile Val Ala Arg Leu Phe Ser Asp  
                   35                  40                  45  
 Ala Arg Arg Leu Leu Leu Tyr Ser Gln Lys Asp Thr Ser Met Lys Asp  
                   50                  55                  60  
 10 Met Arg Lys Val Leu Arg Thr Leu Gln Gln Ile Lys Lys Ser Ser Ser  
                   65                  70                  75                  80  
 Asn Leu Lys Leu Gln Asp Phe Leu Val Asp Asn Glu Thr Phe Ser Gly  
                   85                  90                  95  
 Phe Leu Tyr His Asn Leu Ser Leu Pro Lys Ser Thr Val Asp Lys Met  
                   100                  105                  110  
 20 Leu Arg Ala Asp Val Ile Leu His Lys Val Phe Leu Gln Gly Tyr Gln  
                   115                  120                  125  
 Leu His Leu Thr Ser Leu Cys Asn Gly Ser Lys Ser Glu Glu Met Ile  
                   130                  135                  140  
 25 Gln Leu Gly Asp Gln Glu Val Ser Glu Leu Cys Gly Leu Pro Arg Glu  
                   145                  150                  155                  160  
 Lys Leu Ala Ala Ala Glu Arg Val Leu Arg Ser Asn Met Asp Ile Leu  
                   165                  170                  175  
 30 Lys Pro Ile Leu Arg Thr Leu Asn Ser Thr Ser Pro Phe Pro Ser Lys  
                   180                  185                  190  
 35 Glu Leu Ala Glu Ala Thr Lys Thr Leu Leu His Ser Leu Gly Thr Leu  
                   195                  200                  205  
 Ala Gln Glu Leu Phe Ser Met Arg Ser Trp Ser Asp Met Arg Gln Glu  
                   210                  215                  220  
 40 Val Met Phe Leu Thr Asn Val Asn Ser Ser Ser Ser Ser Thr Gln Ile  
                   225                  230                  235                  240  
 Tyr Gln Ala Val Ser Arg Ile Val Cys Gly His Pro Glu Gly Gly Gly  
                   245                  250                  255  
 45 Leu Lys Ile Lys Ser Leu Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala  
                   260                  265                  270  
 50 Leu Phe Gly Gly Asn Gly Thr Glu Glu Asp Ala Glu Thr Phe Tyr Asp  
                   275                  280                  285  
 Asn Ser Thr Thr Pro Tyr Cys Asn Asp Leu Met Lys Asn Leu Glu Ser  
                   290                  295                  300  
 55 Ser Pro Leu Ser Arg Ile Ile Trp Lys Ala Leu Lys Pro Leu Leu Val  
                   305                  310                  315                  320  
 Gly Lys Ile Leu Tyr Thr Pro Asp Thr Pro Ala Thr Arg Gln Val Met  
                   325                  330                  335  
 60 Ala Glu Val Asn Lys Thr Phe Gln Glu Leu Ala Val Phe His Asp Leu  
                   340                  345                  350  
 65 Glu Gly Met Trp Glu Glu Leu Ser Pro Lys Ile Trp Thr Phe Met Glu  
                   355                  360                  365  
 Asn Ser Gln Glu Met Asp Leu Val Arg Met Leu Leu Asp Ser Arg Asp

	370	375	380	
	Asn Asp His Phe Trp	Glu Gln Gln Leu Asp	Gly Leu Asp Trp Thr Ala	
	385	390	395	400
5	Gln Asp Ile Val Ala	Phe Leu Ala Lys His	Pro Glu Asp Val Gln Ser	
		405	410	415
10	Ser Asn Gly Ser Val Tyr Thr	Trp Arg Glu Ala Phe Asn Glu Thr Asn		
		420	425	430
	Gln Ala Ile Arg Thr Ile Ser	Arg Phe Met Glu Cys Val Asn Leu Asn		
		435	440	445
15	Lys Leu Glu Pro Ile Ala Thr	Glu Val Trp Leu Ile Asn Lys Ser Met		
		450	455	460
	Glu Leu Leu Asp Glu Arg Lys Phe Trp	Ala Gly Ile Val Phe Thr Gly		
		465	470	475
20	Ile Thr Pro Gly Ser Ile Glu Leu Pro	His His Val Lys Tyr Lys Ile		
		485	490	495
25	Arg Met Asp Ile Asp Asn Val Glu	Arg Thr Asn Lys Ile Lys Asp Gly		
		500	505	510
	Tyr Trp Asp Pro Gly Pro Arg	Ala Asp Pro Phe Glu Asp Met Arg Tyr		
		515	520	525
30	Val Trp Gly Gly Phe Ala Tyr	Leu Gln Asp Val Val Glu Gln Ala Ile		
		530	535	540
	Ile Arg Val Leu Thr Gly Thr Glu Lys Lys	Thr Gly Val Tyr Met Gln		
		545	550	555
35	Gln Met Pro Tyr Pro Cys Tyr Val	Asp Asp Ile Phe Leu Arg Val Met		
		565	570	575
40	Ser Arg Ser Met Pro Leu Phe Met	Thr Leu Ala Trp Ile Tyr Ser Val		
		580	585	590
	Ala Val Ile Ile Lys Gly Ile Val Tyr	Glu Lys Glu Ala Arg Leu Lys		
		595	600	605
45	Glu Thr Met Arg Ile Met Gly Leu Asp	Asn Ser Ile Leu Trp Phe Ser		
		610	615	620
	Trp Phe Ile Ser Ser Leu Ile Pro Leu Leu	Val Ser Ala Gly Leu Leu		
		625	630	635
50	Val Val Ile Leu Lys Leu Gly Asn Leu Leu	Pro Tyr Ser Asp Pro Ser		
		645	650	655
	Val Val Phe Val Phe Leu Ser Val Phe	Ala Val Val Thr Ile Leu Gln		
		660	665	670
55	Cys Phe Leu Ile Ser Thr Leu Phe Ser	Arg Ala Asn Leu Ala Ala Ala		
		675	680	685
60	Cys Gly Gly Ile Ile Tyr Phe Thr Leu Tyr	Leu Pro Tyr Val Leu Cys		
		690	695	700
	Val Ala Trp Gln Asp Tyr Val Gly Phe Thr	Leu Lys Ile Phe Ala Ser		
		705	710	715
65	Leu Leu Ser Pro Val Ala Phe Gly Phe	Gly Cys Glu Tyr Phe Ala Leu		
		725	730	735

Phe Glu Glu Gln Gly Ile Gly Val Gln Trp Asp Asn Leu Phe Glu Ser  
 740 745 750  
 5 Pro Val Glu Glu Asp Gly Phe Asn Leu Thr Thr Ser Val Ser Met Met  
 755 760 765  
 Leu Phe Asp Thr Phe Leu Tyr Gly Val Met Thr Trp Tyr Ile Glu Ala  
 770 775 780  
 10 Val Phe Pro Gly Gln Tyr Gly Ile Pro Arg Pro Trp Tyr Phe Pro Cys  
 785 790 795 800  
 Thr Lys Ser Tyr Trp Phe Gly Glu Glu Ser Asp Glu Lys Ser His Pro  
 805 810 815  
 15 Gly Ser Asn Gln Lys Arg Ile Ser Glu Ile Cys Met Glu Glu Glu Pro  
 820 825 830  
 20 Thr His Leu Lys Leu Gly Val Ser Ile Gln Asn Leu Val Lys Val Tyr  
 835 840 845  
 Arg Asp Gly Met Lys Val Ala Val Asp Gly Leu Ala Leu Asn Phe Tyr  
 850 855 860  
 25 Glu Gly Gln Ile Thr Ser Phe Leu Gly His Asn Gly Ala Gly Lys Thr  
 865 870 875 880  
 Thr Thr Met Ser Ile Leu Thr Gly Leu Phe Pro Pro Thr Ser Gly Thr  
 885 890 895  
 30 Ala Tyr Ile Leu Gly Lys Asp Ile Arg Ser Glu Met Ser Thr Ile Arg  
 900 905 910  
 35 Gln Asn Leu Gly Val Cys Pro Gln His Asn Val Leu Phe Asp Met Leu  
 915 920 925  
 Thr Val Glu Glu His Ile Trp Phe Tyr Ala Arg Leu Lys Gly Leu Ser  
 930 935 940  
 40 Glu Lys His Val Lys Ala Glu Met Glu Gln Met Ala Leu Asp Val Gly  
 945 950 955 960  
 Leu Pro Ser Ser Lys Leu Lys Ser Lys Thr Ser Gln Leu Ser Gly Gly  
 965 970 975  
 45 Met Gln Arg Lys Leu Ser Val Ala Leu Ala Phe Val Gly Gly Ser Lys  
 980 985 990  
 50 Val Val Ile Leu Asp Glu Pro Thr Ala Gly Val Asp Pro Tyr Ser Arg  
 995 1000 1005  
 Arg Gly Ile Trp Glu Leu Leu Leu Lys Tyr Arg Gln Gly Arg Thr Ile  
 1010 1015 1020  
 55 Ile Leu Ser Thr His His Met Asp Glu Ala Asp Val Leu Gly Asp Arg  
 1025 1030 1035 1040  
 Ile Ala Ile Ile Ser His Gly Lys Leu Cys Cys Val Gly Ser Ser Leu  
 1045 1050 1055  
 60 Phe Leu Lys Asn Gln Leu Gly Thr Gly Tyr Tyr Leu Thr Leu Val Lys  
 1060 1065 1070  
 65 Lys Asp Val Glu Ser Ser Leu Ser Ser Cys Arg Asn Ser Ser Ser Thr  
 1075 1080 1085  
 Val Ser Tyr Leu Lys Lys Glu Asp Ser Val Ser Gln Ser Ser Ser Asp  
 1090 1095 1100



Ala Gly Leu Gly Ser Asp His Glu Ser Asp Thr Leu Thr Ile Asp Val  
 1105 1110 1115 1120  
 5 Ser Ala Ile Ser Asn Leu Ile Arg Lys His Val Ser Glu Ala Arg Leu  
 1125 1130 1135  
 Val Glu Asp Ile Gly His Glu Leu Thr Tyr Val Leu Pro Tyr Glu Ala  
 1140 1145 1150  
 10 Ala Lys Glu Gly Ala Phe Val Glu Leu Phe His Glu Ile Asp Asp Arg  
 1155 1160 1165  
 Leu Ser Asp Leu Gly Ile Ser Ser Tyr Gly Ile Ser Glu Thr Thr Leu  
 1170 1175 1180  
 15 Glu Glu Ile Phe Leu Lys Val Ala Glu Glu Ser Gly Val Asp Ala Glu  
 1185 1190 1195 1200  
 20 Thr Ser Asp Gly Thr Leu Pro Ala Arg Arg Asn Arg Arg Ala Phe Gly  
 1205 1210 1215  
 Asp Lys Gln Ser Cys Leu Arg Pro Phe Thr Glu Asp Asp Ala Ala Asp  
 1220 1225 1230  
 25 Pro Asn Asp Ser Asp Ile Asp Pro Glu Ser Arg Glu Thr Asp Leu Leu  
 1235 1240 1245  
 Ser Gly Met Asp Gly Lys Gly Ser Tyr Gln Val Lys Gly Trp Lys Leu  
 1250 1255 1260  
 30 Thr Gln Gln Gln Phe Val Ala Leu Leu Trp Lys Arg Leu Leu Ile Ala  
 1265 1270 1275 1280  
 35 Arg Arg Ser Arg Lys Gly Phe Phe Ala Gln Ile Val Leu Pro Ala Val  
 1285 1290 1295  
 Phe Val Cys Ile Ala Leu Val Phe Ser Leu Ile Val Pro Pro Phe Gly  
 1300 1305 1310  
 40 Lys Tyr Pro Ser Leu Glu Leu Gln Pro Trp Met Tyr Asn Glu Gln Tyr  
 1315 1320 1325  
 Thr Phe Val Ser Asn Asp Ala Pro Glu Asp Thr Gly Thr Leu Glu Leu  
 1330 1335 1340  
 Leu Asn Ala Leu Thr Lys Asp Pro Gly Phe Gly Thr Arg Cys Met Glu  
 1345 1350 1355 1360  
 50 Gly Asn Pro Ile Pro Asp Thr Pro Cys Gln Ala Gly Glu Glu Glu Trp  
 1365 1370 1375  
 Thr Thr Ala Pro Val Pro Gln Thr Ile Met Asp Leu Phe Gln Asn Gly  
 1380 1385 1390  
 55 Asn Trp Thr Met Gln Asn Pro Ser Pro Ala Cys Gln Cys Ser Ser Asp  
 1395 1400 1405  
 Lys Ile Lys Lys Met Leu Pro Val Cys Pro Pro Gly Ala Gly Gly Leu  
 1410 1415 1420  
 Pro Pro Pro Gln Arg Lys Gln Asn Thr Ala Asp Ile Leu Gln Asp Leu  
 1425 1430 1435 1440  
 65 Thr Gly Arg Asn Ile Ser Asp Tyr Leu Val Lys Thr Tyr Val Gln Ile  
 1445 1450 1455  
 Ile Ala Lys Ser Leu Lys Asn Lys Ile Trp Val Asn Glu Phe Arg Tyr

	1460	1465	1470
	Gly Gly Phe Ser Leu Gly Val Ser Asn Thr Gln Ala Leu Pro Pro Ser		
	1475	1480	1485
5	Gln Glu Val Asn Asp Ala Thr Lys Gln Met Lys Lys His Leu Lys Leu		
	1490	1495	1500
	Ala Lys Asp Ser Ser Ala Asp Arg Phe Leu Asn Ser Leu Gly Arg Phe		
10	1505	1510	1515 1520
	Met Thr Gly Leu Asp Thr Arg Asn Asn Val Lys Val Trp Phe Asn Asn		
	1525	1530	1535
15	Lys Gly Trp His Ala Ile Ser Ser Phe Leu Asn Val Ile Asn Asn Ala		
	1540	1545	1550
	Ile Leu Arg Ala Asn Leu Gln Lys Gly Glu Asn Pro Ser His Tyr Gly		
20	1555	1560	1565
	Ile Thr Ala Phe Asn His Pro Leu Asn Leu Thr Lys Gln Gln Leu Ser		
	1570	1575	1580
	Glu Val Ala Pro Met Thr Thr Ser Val Asp Val Leu Val Ser Ile Cys		
25	1585	1590	1595 1600
	Val Ile Phe Ala Met Ser Phe Val Pro Ala Ser Phe Val Val Phe Leu		
	1605	1610	1615
30	Ile Gln Glu Arg Val Ser Lys Ala Lys His Leu Gln Phe Ile Ser Gly		
	1620	1625	1630
	Val Lys Pro Val Ile Tyr Trp Leu Ser Asn Phe Val Trp Asp Met Cys		
35	1635	1640	1645
	Asn Tyr Val Val Pro Ala Thr Leu Val Ile Ile Ile Phe Ile Cys Phe		
	1650	1655	1660
	Gln Gln Lys Ser Tyr Val Ser Ser Thr Asn Leu Pro Val Leu Ala Leu		
40	1665	1670	1675 1680
	Leu Leu Leu Leu Tyr Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala		
	1685	1690	1695
45	Ser Phe Val Phe Lys Ile Pro Ser Thr Ala Tyr Val Val Leu Thr Ser		
	1700	1705	1710
	Val Asn Leu Phe Ile Gly Ile Asn Gly Ser Val Ala Thr Phe Val Leu		
50	1715	1720	1725
	Glu Leu Phe Thr Asp Asn Lys Leu Asn Asn Ile Asn Asp Ile Leu Lys		
	1730	1735	1740
	Ser Val Phe Leu Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile		
55	1745	1750	1755 1760
	Asp Met Val Lys Asn Gln Ala Met Ala Asp Ala Leu Glu Arg Phe Gly		
	1765	1770	1775
60	Glu Asn Arg Phe Val Ser Pro Leu Ser Trp Asp Leu Val Gly Arg Asn		
	1780	1785	1790
	Leu Phe Ala Met Ala Val Glu Gly Val Val Phe Phe Leu Ile Thr Val		
65	1795	1800	1805
	Leu Ile Gln Tyr Arg Phe Phe Ile Arg Pro Arg Pro Val Asn Ala Lys		
	1810	1815	1820

Leu Ser Pro Leu Asn Asp Glu Asp Glu Asp Val Arg Arg Glu Arg Gln  
 1825 1830 1835 1840  
 Arg Ile Leu Asp Gly Gly Gly Gln Asn Asp Ile Leu Glu Ile Lys Glu  
 5 1845 1850 1855  
 Leu Thr Lys Ile Tyr Arg Arg Lys Arg Lys Pro Ala Val Asp Arg Ile  
 1860 1865 1870  
 10 Cys Val Gly Ile Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn  
 1875 1880 1885  
 Gly Ala Gly Lys Ser Ser Thr Phe Lys Met Leu Thr Gly Asp Thr Thr  
 1890 1895 1900  
 15 Val Thr Arg Gly Asp Ala Phe Leu Asn Arg Asn Ser Ile Leu Ser Asn  
 1905 1910 1915 1920  
 Ile His Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala  
 20 1925 1930 1935  
 Ile Thr Glu Leu Leu Thr Gly Arg Glu His Val Glu Phe Phe Ala Leu  
 1940 1945 1950  
 25 Leu Arg Gly Val Pro Glu Lys Glu Val Gly Lys Val Gly Glu Trp Ala  
 1955 1960 1965  
 Ile Arg Lys Leu Gly Leu Val Lys Tyr Gly Glu Lys Tyr Ala Gly Asn  
 1970 1975 1980  
 30 Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala Met Ala Leu Ile  
 1985 1990 1995 2000  
 Gly Gly Pro Pro Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp  
 35 2005 2010 2015  
 Pro Lys Ala Arg Arg Phe Leu Trp Asn Cys Ala Leu Ser Val Val Lys  
 2020 2025 2030  
 40 Glu Gly Arg Ser Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu  
 2035 2040 2045  
 Ala Leu Cys Thr Arg Met Ala Ile Met Val Asn Gly Arg Phe Arg Cys  
 2050 2055 2060  
 45 Leu Gly Ser Val Gln His Leu Lys Asn Arg Phe Gly Asp Gly Tyr Thr  
 2065 2070 2075 2080  
 Ile Val Val Arg Ile Ala Gly Ser Asn Pro Asp Leu Lys Pro Val Gln  
 50 2085 2090 2095  
 Asp Phe Phe Gly Leu Ala Phe Pro Gly Ser Val Pro Lys Glu Lys His  
 2100 2105 2110  
 55 Arg Asn Met Leu Gln Tyr Gln Leu Pro Ser Ser Leu Ser Ser Leu Ala  
 2115 2120 2125  
 Arg Ile Phe Ser Ile Leu Ser Gln Ser Lys Lys Arg Leu His Ile Glu  
 2130 2135 2140  
 60 Asp Tyr Ser Val Ser Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe  
 2145 2150 2155 2160  
 Ala Lys Asp Gln Ser Asp Asp Asp His Leu Lys Asp Leu Ser Leu His  
 65 2165 2170 2175  
 Lys Asn Gln Thr Val Val Asp Val Ala Val Leu Thr Ser Phe Leu Gln  
 2180 2185 2190

Asp Glu Lys Val Lys Glu Ser Tyr Val  
2195 2200

5

&lt;210&gt; 140

&lt;211&gt; 2233

&lt;212&gt; PRT

10 &lt;213&gt; Homo sapiens

&lt;400&gt; 140

Met Pro Ser Ala Gly Thr Leu Pro Trp Val Gln Gly Ile Ile Cys Asn  
1 5 10 15

15

Ala Asn Asn Pro Cys Phe Arg Tyr Pro Thr Pro Gly Glu Ala Pro Gly  
20 25 30

20

Val Val Gly Asn Phe Asn Lys Ser Ile Val Ala Arg Leu Phe Ser Asp  
35 40 45

Ala Arg Arg Leu Leu Leu Tyr Ser Gln Lys Asp Thr Ser Met Lys Asp  
50 55 60

25

Met Arg Lys Val Leu Arg Thr Leu Gln Gln Ile Lys Lys Ser Ser Ser  
65 70 75 80

Asn Leu Lys Leu Gln Asp Phe Leu Val Asp Asn Glu Thr Phe Ser Gly  
85 90 95

30

Phe Leu Tyr His Asn Leu Ser Leu Pro Lys Ser Thr Val Asp Lys Met  
100 105 110

35

Leu Arg Ala Asp Val Ile Leu His Lys Val Phe Leu Gln Gly Tyr Gln  
115 120 125

Leu His Leu Thr Ser Leu Cys Asn Gly Ser Lys Ser Glu Glu Met Ile  
130 135 140

40

Gln Leu Gly Asp Gln Glu Val Ser Glu Leu Cys Gly Leu Pro Arg Glu  
145 150 155 160

Lys Leu Ala Ala Ala Glu Arg Val Leu Arg Ser Asn Met Asp Ile Leu  
165 170 175

45

Lys Pro Ile Leu Arg Thr Leu Asn Ser Thr Ser Pro Phe Pro Ser Lys  
180 185 190

50

Glu Leu Ala Glu Ala Thr Lys Thr Leu Leu His Ser Leu Gly Thr Leu  
195 200 205

Ala Gln Glu Leu Phe Ser Met Arg Ser Trp Ser Asp Met Arg Gln Glu  
210 215 220

55

Val Met Phe Leu Thr Asn Val Asn Ser Ser Ser Ser Ser Thr Gln Ile  
225 230 235 240

Tyr Gln Ala Val Ser Arg Ile Val Cys Gly His Pro Glu Gly Gly Gly  
245 250 255

60

Leu Lys Ile Lys Ser Leu Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala  
260 265 270

65

Leu Phe Gly Gly Asn Gly Thr Glu Glu Asp Ala Glu Thr Phe Tyr Asp  
275 280 285

Asn Ser Thr Thr Pro Tyr Cys Asn Asp Leu Met Lys Asn Leu Glu Ser  
290 295 300

Ser Pro Leu Ser Arg Ile Ile Trp Lys Ala Leu Lys Pro Leu Leu Val  
 305 310 315 320  
 5 Gly Lys Ile Leu Tyr Thr Pro Asp Thr Pro Ala Thr Arg Gln Val Met  
 325 330 335  
 Ala Glu Val Asn Lys Thr Phe Gln Glu Leu Ala Val Phe His Asp Leu  
 340 345 350  
 10 Glu Gly Met Trp Glu Glu Leu Ser Pro Lys Ile Trp Thr Phe Met Glu  
 355 360 365  
 Asn Ser Gln Glu Met Asp Leu Val Arg Met Leu Leu Asp Ser Arg Asp  
 15 370 375 380  
 Asn Asp His Phe Trp Glu Gln Gln Leu Asp Gly Leu Asp Trp Thr Ala  
 385 390 395 400  
 20 Gln Asp Ile Val Ala Phe Leu Ala Lys His Pro Glu Asp Val Gln Ser  
 405 410 415  
 Ser Asn Gly Ser Val Tyr Thr Trp Arg Glu Ala Phe Asn Glu Thr Asn  
 420 425 430  
 25 Gln Ala Ile Arg Thr Ile Ser Arg Phe Met Glu Cys Val Asn Leu Asn  
 435 440 445  
 Lys Leu Glu Pro Ile Ala Thr Glu Val Trp Leu Ile Asn Lys Ser Met  
 30 450 455 460  
 Glu Leu Leu Glu Tyr Ser Gly Val Thr Ser Ala His Cys Asn Leu Cys  
 465 470 475 480  
 35 Leu Leu Ser Ser Ser Asp Ser Arg Ala Ser Ala Ser Gln Val Ala Gly  
 485 490 495  
 Ile Thr Ala Pro Ala Thr Thr Pro Gly Ala Gly Ile Val Phe Thr Gly  
 500 505 510  
 40 Ile Thr Pro Gly Ser Ile Glu Leu Pro His His Val Lys Tyr Lys Ile  
 515 520 525  
 Arg Met Asp Ile Asp Asn Val Glu Arg Thr Asn Lys Ile Lys Asp Gly  
 45 530 535 540  
 Tyr Trp Asp Pro Gly Pro Arg Ala Asp Pro Phe Glu Asp Met Arg Tyr  
 545 550 555 560  
 50 Val Trp Gly Gly Phe Ala Tyr Leu Gln Asp Val Val Glu Gln Ala Ile  
 565 570 575  
 Ile Arg Val Leu Thr Gly Thr Glu Lys Lys Thr Gly Val Tyr Met Gln  
 580 585 590  
 55 Gln Met Pro Tyr Pro Cys Tyr Val Asp Asp Ile Phe Leu Arg Val Met  
 595 600 605  
 Ser Arg Ser Met Pro Leu Phe Met Thr Leu Ala Trp Ile Tyr Ser Val  
 60 610 615 620  
 Ala Val Ile Ile Lys Gly Ile Val Tyr Glu Lys Glu Ala Arg Leu Lys  
 625 630 635 640  
 65 Glu Thr Met Arg Ile Met Gly Leu Asp Asn Ser Ile Leu Trp Phe Ser  
 645 650 655  
 Trp Phe Ile Ser Ser Leu Ile Pro Leu Leu Val Ser Ala Gly Leu Leu

	660							665							670						
	Val	Val	Ile	Leu	Lys	Leu	Gly	Asn	Leu	Leu	Pro	Tyr	Ser	Asp	Pro	Ser					
	675							680							685						
5	Val	Val	Phe	Val	Phe	Leu	Ser	Val	Phe	Ala	Val	Val	Thr	Ile	Leu	Gln					
	690							695							700						
10	Cys	Phe	Leu	Ile	Ser	Thr	Leu	Phe	Ser	Arg	Ala	Asn	Leu	Ala	Ala	Ala					
	705							710							715						
	Cys	Gly	Gly	Ile	Ile	Tyr	Phe	Thr	Leu	Tyr	Leu	Pro	Tyr	Val	Leu	Cys					
	725							730							735						
15	Val	Ala	Trp	Gln	Asp	Tyr	Val	Gly	Phe	Thr	Leu	Lys	Ile	Phe	Ala	Ser					
	740							745							750						
	Leu	Leu	Ser	Pro	Val	Ala	Phe	Gly	Phe	Gly	Cys	Glu	Tyr	Phe	Ala	Leu					
	755							760							765						
20	Phe	Glu	Glu	Gln	Gly	Ile	Gly	Val	Gln	Trp	Asp	Asn	Leu	Phe	Glu	Ser					
	770							775							780						
25	Pro	Val	Glu	Glu	Asp	Gly	Phe	Asn	Leu	Thr	Thr	Ser	Val	Ser	Met	Met					
	785							790							795						
	Leu	Phe	Asp	Thr	Phe	Leu	Tyr	Gly	Val	Met	Thr	Trp	Tyr	Ile	Glu	Ala					
	805							810							815						
30	Val	Phe	Pro	Gly	Gln	Tyr	Gly	Ile	Pro	Arg	Pro	Trp	Tyr	Phe	Pro	Cys					
	820							825							830						
	Thr	Lys	Ser	Tyr	Trp	Phe	Gly	Glu	Glu	Ser	Asp	Glu	Lys	Ser	His	Pro					
	835							840							845						
35	Gly	Ser	Asn	Gln	Lys	Arg	Ile	Ser	Glu	Ile	Cys	Met	Glu	Glu	Glu	Pro					
	850							855							860						
40	Thr	His	Leu	Lys	Leu	Gly	Val	Ser	Ile	Gln	Asn	Leu	Val	Lys	Val	Tyr					
	865							870							875						
	Arg	Asp	Gly	Met	Lys	Val	Ala	Val	Asp	Gly	Leu	Ala	Leu	Asn	Phe	Tyr					
	885							890							895						
45	Glu	Gly	Gln	Ile	Thr	Ser	Phe	Leu	Gly	His	Asn	Gly	Ala	Gly	Lys	Thr					
	900							905							910						
	Thr	Thr	Met	Ser	Ile	Leu	Thr	Gly	Leu	Phe	Pro	Pro	Thr	Ser	Gly	Thr					
	915							920							925						
50	Ala	Tyr	Ile	Leu	Gly	Lys	Asp	Ile	Arg	Ser	Glu	Met	Ser	Thr	Ile	Arg					
	930							935							940						
55	Gln	Asn	Leu	Gly	Val	Cys	Pro	Gln	His	Asn	Val	Leu	Phe	Asp	Met	Leu					
	945							950							955						
	Thr	Val	Glu	Glu	His	Ile	Trp	Phe	Tyr	Ala	Arg	Leu	Lys	Gly	Leu	Ser					
	965							970							975						
60	Glu	Lys	His	Val	Lys	Ala	Glu	Met	Glu	Gln	Met	Ala	Leu	Asp	Val	Gly					
	980							985							990						
	Leu	Pro	Ser	Ser	Lys	Leu	Lys	Ser	Lys	Thr	Ser	Gln	Leu	Ser	Gly	Gly					
	995							1000							1005						
65	Met	Gln	Arg	Lys	Leu	Ser	Val	Ala	Leu	Ala	Phe	Val	Gly	Gly	Ser	Lys					
	1010							1015							1020						

Val Val Ile Leu Asp Glu Pro Thr Ala Gly Val Asp Pro Tyr Ser Arg  
 1025 1030 1035 1040  
 5 Arg Gly Ile Trp Glu Leu Leu Leu Lys Tyr Arg Gln Gly Arg Thr Ile  
 1045 1050 1055  
 Ile Leu Ser Thr His His Met Asp Glu Ala Asp Val Leu Gly Asp Arg  
 1060 1065 1070  
 10 Ile Ala Ile Ile Ser His Gly Lys Leu Cys Cys Val Gly Ser Ser Leu  
 1075 1080 1085  
 Phe Leu Lys Asn Gln Leu Gly Thr Gly Tyr Tyr Leu Thr Leu Val Lys  
 1090 1095 1100  
 15 Lys Asp Val Glu Ser Ser Leu Ser Ser Cys Arg Asn Ser Ser Ser Thr  
 1105 1110 1115 1120  
 Val Ser Tyr Leu Lys Lys Glu Asp Ser Val Ser Gln Ser Ser Ser Asp  
 1125 1130 1135  
 Ala Gly Leu Gly Ser Asp His Glu Ser Asp Thr Leu Thr Ile Asp Val  
 1140 1145 1150  
 25 Ser Ala Ile Ser Asn Leu Ile Arg Lys His Val Ser Glu Ala Arg Leu  
 1155 1160 1165  
 Val Glu Asp Ile Gly His Glu Leu Thr Tyr Val Leu Pro Tyr Glu Ala  
 1170 1175 1180  
 30 Ala Lys Glu Gly Ala Phe Val Glu Leu Phe His Glu Ile Asp Asp Arg  
 1185 1190 1195 1200  
 Leu Ser Asp Leu Gly Ile Ser Ser Tyr Gly Ile Ser Glu Thr Thr Leu  
 1205 1210 1215  
 Glu Glu Ile Phe Leu Lys Val Ala Glu Glu Ser Gly Val Asp Ala Glu  
 1220 1225 1230  
 40 Thr Ser Asp Gly Thr Leu Pro Ala Arg Arg Asn Arg Arg Ala Phe Gly  
 1235 1240 1245  
 Asp Lys Gln Ser Cys Leu Arg Pro Phe Thr Glu Asp Asp Ala Ala Asp  
 1250 1255 1260  
 45 Pro Asn Asp Ser Asp Ile Asp Pro Glu Ser Arg Glu Thr Asp Leu Leu  
 1265 1270 1275 1280  
 Ser Gly Met Asp Gly Lys Gly Ser Tyr Gln Val Lys Gly Trp Lys Leu  
 1285 1290 1295  
 50 Thr Gln Gln Gln Phe Val Ala Leu Leu Trp Lys Arg Leu Leu Ile Ala  
 1300 1305 1310  
 Arg Arg Ser Arg Lys Gly Phe Phe Ala Gln Ile Val Leu Pro Ala Val  
 1315 1320 1325  
 Phe Val Cys Ile Ala Leu Val Phe Ser Leu Ile Val Pro Pro Phe Gly  
 1330 1335 1340  
 60 Lys Tyr Pro Ser Leu Glu Leu Gln Pro Trp Met Tyr Asn Glu Gln Tyr  
 1345 1350 1355 1360  
 Thr Phe Val Ser Asn Asp Ala Pro Glu Asp Thr Gly Thr Leu Glu Leu  
 1365 1370 1375  
 65 Leu Asn Ala Leu Thr Lys Asp Pro Gly Phe Gly Thr Arg Cys Met Glu  
 1380 1385 1390

Gly Asn Pro Ile Pro Asp Thr Pro Cys Gln Ala Gly Glu Glu Glu Trp  
 1395 1400 1405  
 5 Thr Thr Ala Pro Val Pro Gln Thr Ile Met Asp Leu Phe Gln Asn Gly  
 1410 1415 1420  
 Asn Trp Thr Met Gln Asn Pro Ser Pro Ala Cys Gln Cys Ser Ser Asp  
 1425 1430 1435 1440  
 10 Lys Ile Lys Lys Met Leu Pro Val Cys Pro Pro Gly Ala Gly Gly Leu  
 1445 1450 1455  
 Pro Pro Pro Gln Arg Lys Gln Asn Thr Ala Asp Ile Leu Gln Asp Leu  
 1460 1465 1470  
 15 Thr Gly Arg Asn Ile Ser Asp Tyr Leu Val Lys Thr Tyr Val Gln Ile  
 1475 1480 1485  
 20 Ile Ala Lys Ser Leu Lys Asn Lys Ile Trp Val Asn Glu Phe Arg Tyr  
 1490 1495 1500  
 Gly Gly Phe Ser Leu Gly Val Ser Asn Thr Gln Ala Leu Pro Pro Ser  
 1505 1510 1515 1520  
 25 Gln Glu Val Asn Asp Ala Thr Lys Gln Met Lys Lys His Leu Lys Leu  
 1525 1530 1535  
 Ala Lys Asp Ser Ser Ala Asp Arg Phe Leu Asn Ser Leu Gly Arg Phe  
 1540 1545 1550  
 30 Met Thr Gly Leu Asp Thr Arg Asn Asn Val Lys Val Trp Phe Asn Asn  
 1555 1560 1565  
 35 Lys Gly Trp His Ala Ile Ser Ser Phe Leu Asn Val Ile Asn Asn Ala  
 1570 1575 1580  
 Ile Leu Arg Ala Asn Leu Gln Lys Gly Glu Asn Pro Ser His Tyr Gly  
 1585 1590 1595 1600  
 40 Ile Thr Ala Phe Asn His Pro Leu Asn Leu Thr Lys Gln Gln Leu Ser  
 1605 1610 1615  
 Glu Val Ala Pro Met Thr Thr Ser Val Asp Val Leu Val Ser Ile Cys  
 1620 1625 1630  
 45 Val Ile Phe Ala Met Ser Phe Val Pro Ala Ser Phe Val Val Phe Leu  
 1635 1640 1645  
 50 Ile Gln Glu Arg Val Ser Lys Ala Lys His Leu Gln Phe Ile Ser Gly  
 1650 1655 1660  
 Val Lys Pro Val Ile Tyr Trp Leu Ser Asn Phe Val Trp Asp Met Cys  
 1665 1670 1675 1680  
 55 Asn Tyr Val Val Pro Ala Thr Leu Val Ile Ile Ile Phe Ile Cys Phe  
 1685 1690 1695  
 Gln Gln Lys Ser Tyr Val Ser Ser Thr Asn Leu Pro Val Leu Ala Leu  
 1700 1705 1710  
 60 Leu Leu Leu Leu Tyr Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala  
 1715 1720 1725  
 65 Ser Phe Val Phe Lys Ile Pro Ser Thr Ala Tyr Val Val Leu Thr Ser  
 1730 1735 1740  
 Val Asn Leu Phe Ile Gly Ile Asn Gly Ser Val Ala Thr Phe Val Leu



	1745	1750	1755	1760
	Glu Leu Phe Thr Asp Asn Lys Leu Asn Asn Ile Asn Asp Ile Leu Lys			
	1765		1770	1775
5	Ser Val Phe Leu Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile			
	1780	1785		1790
10	Asp Met Val Lys Asn Gln Ala Met Ala Asp Ala Leu Glu Arg Phe Gly			
	1795	1800		1805
	Glu Asn Arg Phe Val Ser Pro Leu Ser Trp Asp Leu Val Gly Arg Asn			
	1810	1815		1820
15	Leu Phe Ala Met Ala Val Glu Gly Val Val Phe Phe Leu Ile Thr Val			
	1825	1830	1835	1840
	Leu Ile Gln Tyr Arg Phe Phe Ile Arg Pro Arg Pro Val Asn Ala Lys			
	1845	1850		1855
20	Leu Ser Pro Leu Asn Asp Glu Asp Glu Asp Val Arg Arg Glu Arg Gln			
	1860	1865		1870
25	Arg Ile Leu Asp Gly Gly Gly Gln Asn Asp Ile Leu Glu Ile Lys Glu			
	1875	1880		1885
	Leu Thr Lys Ile Tyr Arg Arg Lys Arg Lys Pro Ala Val Asp Arg Ile			
	1890	1895		1900
30	Cys Val Gly Ile Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn			
	1905	1910	1915	1920
	Gly Ala Gly Lys Ser Ser Thr Phe Lys Met Leu Thr Gly Asp Thr Thr			
	1925	1930		1935
35	Val Thr Arg Gly Asp Ala Phe Leu Asn Arg Asn Ser Ile Leu Ser Asn			
	1940	1945		1950
40	Ile His Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala			
	1955	1960		1965
	Ile Thr Glu Leu Leu Thr Gly Arg Glu His Val Glu Phe Phe Ala Leu			
	1970	1975		1980
45	Leu Arg Gly Val Pro Glu Lys Glu Val Gly Lys Val Gly Glu Trp Ala			
	1985	1990	1995	2000
	Ile Arg Lys Leu Gly Leu Val Lys Tyr Gly Glu Lys Tyr Ala Gly Asn			
	2005	2010		2015
50	Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala Met Ala Leu Ile			
	2020	2025		2030
55	Gly Gly Pro Pro Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp			
	2035	2040		2045
	Pro Lys Ala Arg Arg Phe Leu Trp Asn Cys Ala Leu Ser Val Val Lys			
	2050	2055		2060
60	Glu Gly Arg Ser Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu			
	2065	2070	2075	2080
	Ala Leu Cys Thr Arg Met Ala Ile Met Val Asn Gly Arg Phe Arg Cys			
	2085	2090		2095
65	Leu Gly Ser Val Gln His Leu Lys Asn Arg Phe Gly Asp Gly Tyr Thr			
	2100	2105		2110

Ile Val Val Arg Ile Ala Gly Ser Asn Pro Asp Leu Lys Pro Val Gln  
 2115 2120 2125  
 5 Asp Phe Phe Gly Leu Ala Phe Pro Gly Ser Val Pro Lys Glu Lys His  
 2130 2135 2140  
 Arg Asn Met Leu Gln Tyr Gln Leu Pro Ser Ser Leu Ser Ser Leu Ala  
 2145 2150 2155 2160  
 10 Arg Ile Phe Ser Ile Leu Ser Gln Ser Lys Lys Arg Leu His Ile Glu  
 2165 2170 2175  
 Asp Tyr Ser Val Ser Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe  
 2180 2185 2190  
 15 Ala Lys Asp Gln Ser Asp Asp Asp His Leu Lys Asp Leu Ser Leu His  
 2195 2200 2205  
 Lys Asn Gln Thr Val Val Asp Val Ala Val Leu Thr Ser Phe Leu Gln  
 2210 2215 2220  
 Asp Glu Lys Val Lys Glu Ser Tyr Val  
 2225 2230  
 25  
 <210> 141  
 <211> 574  
 <212> PRT  
 30 <213> Homo sapiens  
 <400> 141  
 Met Pro Ser Ala Gly Thr Leu Pro Trp Val Gln Gly Ile Ile Cys Asn  
 1 5 10 15  
 35 Ala Asn Asn Pro Cys Phe Arg Tyr Pro Thr Pro Gly Glu Ala Pro Gly  
 20 25 30  
 Val Val Gly Asn Phe Asn Lys Ser Ile Val Ala Arg Leu Phe Ser Asp  
 35 40 45  
 Ala Arg Arg Leu Leu Leu Tyr Ser Gln Lys Asp Thr Ser Met Lys Asp  
 50 55 60  
 45 Met Arg Lys Val Leu Arg Thr Leu Gln Gln Ile Lys Lys Ser Ser Ser  
 65 70 75 80  
 Asn Leu Lys Leu Gln Asp Phe Leu Val Asp Asn Glu Thr Phe Ser Gly  
 85 90 95  
 50 Phe Leu Tyr His Asn Leu Ser Leu Pro Lys Ser Thr Val Asp Lys Met  
 100 105 110  
 Leu Arg Ala Asp Val Ile Leu His Lys Val Phe Leu Gln Gly Tyr Gln  
 115 120 125  
 55 Leu His Leu Thr Ser Leu Cys Asn Gly Ser Lys Ser Glu Glu Met Ile  
 130 135 140  
 60 Gln Leu Gly Asp Gln Glu Val Ser Glu Leu Cys Gly Leu Pro Arg Glu  
 145 150 155 160  
 Lys Leu Ala Ala Ala Glu Arg Val Leu Arg Ser Asn Met Asp Ile Leu  
 165 170 175  
 65 Lys Pro Ile Leu Arg Thr Leu Asn Ser Thr Ser Pro Phe Pro Ser Lys  
 180 185 190

Glu Leu Ala Glu Ala Thr Lys Thr Leu Leu His Ser Leu Gly Thr Leu  
 195 200 205  
 5 Ala Gln Glu Leu Phe Ser Met Arg Ser Trp Ser Asp Met Arg Gln Glu  
 210 215 220  
 Val Met Phe Leu Thr Asn Val Asn Ser Ser Ser Ser Ser Thr Gln Ile  
 225 230 235 240  
 10 Tyr Gln Ala Val Ser Arg Ile Val Cys Gly His Pro Glu Gly Gly Gly  
 245 250 255  
 Leu Lys Ile Lys Ser Leu Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala  
 260 265 270  
 15 Leu Phe Gly Gly Asn Gly Thr Glu Glu Asp Ala Glu Thr Phe Tyr Asp  
 275 280 285  
 Asn Ser Thr Thr Pro Tyr Cys Asn Asp Leu Met Lys Asn Leu Glu Ser  
 290 295 300  
 Ser Pro Leu Ser Arg Ile Ile Trp Lys Ala Leu Lys Pro Leu Leu Val  
 305 310 315 320  
 25 Gly Lys Ile Leu Tyr Thr Pro Asp Thr Pro Ala Thr Arg Gln Val Met  
 325 330 335  
 Ala Glu Val Asn Lys Thr Phe Gln Glu Leu Ala Val Phe His Asp Leu  
 340 345 350  
 30 Glu Gly Met Trp Glu Glu Leu Ser Pro Lys Ile Trp Thr Phe Met Glu  
 355 360 365  
 Asn Ser Gln Glu Met Asp Leu Val Arg Met Leu Leu Asp Ser Arg Asp  
 370 375 380  
 Asn Asp His Phe Trp Glu Gln Gln Leu Asp Gly Leu Asp Trp Thr Ala  
 385 390 395 400  
 40 Gln Asp Ile Val Ala Phe Leu Ala Lys His Pro Glu Asp Val Gln Ser  
 405 410 415  
 Ser Asn Gly Ser Val Tyr Thr Trp Arg Glu Ala Phe Asn Glu Thr Asn  
 420 425 430  
 45 Gln Ala Ile Arg Thr Ile Ser Arg Phe Met Glu Cys Val Asn Leu Asn  
 435 440 445  
 Lys Leu Glu Pro Ile Ala Thr Glu Val Trp Leu Ile Asn Lys Ser Met  
 450 455 460  
 Glu Leu Leu Asp Glu Arg Lys Phe Trp Ala Gly Ile Val Phe Thr Gly  
 465 470 475 480  
 55 Ile Thr Pro Gly Ser Ile Glu Leu Pro His His Val Lys Tyr Lys Ile  
 485 490 495  
 Arg Met Asp Ile Asp Asn Val Glu Arg Thr Asn Lys Ile Lys Asp Gly  
 500 505 510  
 60 Tyr Trp Asp Pro Gly Pro Arg Ala Asp Pro Phe Glu Asp Met Arg Tyr  
 515 520 525  
 Val Trp Gly Gly Phe Ala Tyr Leu Gln Asp Val Val Glu Gln Ala Ile  
 530 535 540  
 65 Ile Arg Val Leu Arg Ala Pro Arg Arg Lys Leu Val Ser Ile Cys Asn  
 545 550 555 560

Arg Cys Pro Ile Pro Val Thr Leu Met Thr Ser Phe Cys Gly  
565 570

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 00/01595

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/52 C07K14/705 C12Q1/68 C07K16/18

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

STRAND, EPO-Internal, BIOSIS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>LANGMANN T, KLUCKEN J, REIL M, LIEBISCH G, LUCIANI MF, CHIMINI G, KAMINSKI WE, SCHMITZ G: "Molecular cloning of the human ATP-binding cassette transporter 1 (hABC1): evidence for sterol-dependent regulation in macrophages" BIOCHEM BIOPHYS RES COMMUN, vol. 257, no. 1, 2 April 1999 (1999-04-02), pages 29-33, XP000877240 cited in the application the whole document</p> <p>---</p> <p>-/--</p>	2-5,14, 22-31

☒ Further documents are listed in the continuation of box C.☐ Patent family members are listed in annex.

## \* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Z" document member of the same patent family

Date of the actual completion of the international search

13 October 2000

Date of mailing of the international search report

23.11.00

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax: (+31-70) 340-3015

Authorized officer

CHAMBONNET, F

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/FR 00/01595

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>EMBASE : AC AQ061641; Cr��e le 03-AUG-1998 (Rel. 56, Created) Caract��risation de la s��quence : "CIT-HSP-2348011.TF CIT-HSP Homo sapiens genomic clone 2348011, Genomic survey sequence.Homo sapiens (human)" XP002132840 the whole document</p> <p>---</p>	1-5
A	<p>RUST, S. ET AL.: "Assignment of Tangier disease to chromosome 9q31 by a graphical linkage exclusion strategy." NAT GENET., vol. 20, no. 1, September 1998 (1998-09), pages 96-98, XP000884511 cited in the application the whole document</p> <p>---</p>	1
A	<p>LUCIANI M F ET AL: "CLONING OF TWO NOVEL ABC TRANSPORTERS MAPPING ON HUMAN CHROMOSOME 9" GENOMICS,US,ACADEMIC PRESS, SAN DIEGO, vol. 21, no. 1, 1 May 1994 (1994-05-01), pages 150-159, XP000869719 ISSN: 0888-7543 cited in the application the whole document</p> <p>---</p>	1
A	<p>LISCUM L, MUNN NJ.: "Intracellular cholesterol transport." BIOCHIM BIOPHYS ACTA., vol. 1438, no. 1, 19 April 1999 (1999-04-19), pages 19-37, XP000889761 page 33, column 1, paragraph 5.2 -column 2</p> <p>---</p>	1
P,X	<p>GURA, T.: "Gene linked to faulty cholesterol transport." SCIENCE., vol. 285, no. 5429, 6 August 1999 (1999-08-06), pages 814-815, XP000877245 the whole document</p> <p>---</p>	1
P,X	<p>RUST S, ROSIER M, FUNKE H, REAL J, AMOURA Z, PIETTE JC, DELEUZE JF, BREWER HB, DUVERGER N, DENEFFLE P, ASSMANN G.: "Tangier disease is caused by mutations in the gene encoding ATP-binding cassette transporter 1." NAT GENET., vol. 22, no. 4, August 1999 (1999-08), pages 352-355, XP000884993 the whole document</p> <p>---</p>	1-5

-/--

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 00/01595

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	REMALEY AT, ET AL.: "Human ATP-binding cassette transporter 1 (ABC1): genomic organization and identification of the genetic defect in the original Tangier disease kindred." PROC NATL ACAD SCI U S A., vol. 96, no. 22, 26 October 1999 (1999-10-26), pages 12685-12690, XP00087247 the whole document	1-5
P,X	--- LAWN RM, WADE DP, GARVIN MR, WANG X, SCHWARTZ K, PORTER JG, SEILHAMER JJ, VAUGHAN AM, ORAM JF.: "The Tangier disease gene product ABC1 controls the cellular apolipoprotein-mediated lipid removal pathway." J CLIN INVEST., vol. 104, no. 8, October 1999 (1999-10), pages R25-R31, XP000884782 the whole document	1-5
P,X	--- BODZIOCH, M. ET AL.: "The gene encoding ATP-binding cassette transporter 1 is mutated in Tangier disease." NAT GENET., vol. 22, no. 4, August 1999 (1999-08), pages 347-351, XP000889766 the whole document	1
P,X	--- MARCIL M, ET AL.: "Mutations in the ABC1 gene in familial HDL deficiency with defective cholesterol efflux." LANCET, vol. 354, no. 9187, 16 October 1999 (1999-10-16), pages 1341-1346, XP00087242 the whole document	1
P,X	--- ORSO, E. ET AL.: "Transport of lipids from golgi to plasma membrane is defective in Tangier disease patients and Abcl-deficient mice." NAT GENET. 2000 FEB;24(2):192-6., XP000889762 the whole document	1
P,X	--- BROOKS-WILSON A, ET AL.: "Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency." NAT GENET., vol. 22, no. 4, August 1999 (1999-08), pages 336-345, XP000889767 the whole document -----	1

# RAPPORT DE RECHERCHE INTERNATIONALE

mande Internationale No  
PCT/FR 00/01595

<b>A. CLASSEMENT DE L'OBJET DE LA DEMANDE</b> CIB 7 C12N15/52 C07K14/705 C12Q1/68 C07K16/18		
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB		
<b>B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE</b> Documentation minimale consultée (système de classification suivi des symboles de classement) CIB 7 C07K		
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche		
Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés) STRAND, EPO-Internal, BIOSIS		
<b>C. DOCUMENTS CONSIDERES COMME PERTINENTS</b>		
Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	LANGMANN T, KLUCKEN J, REIL M, LIEBISCH G, LUCIANI MF, CHIMINI G, KAMINSKI WE, SCHMITZ G: "Molecular cloning of the human ATP-binding cassette transporter 1 (hABC1): evidence for sterol-dependent regulation in macrophages" BIOCHEM BIOPHYS RES COMMUN, vol. 257, no. 1, 2 avril 1999 (1999-04-02), pages 29-33, XP000877240 cité dans la demande le document en entier --- -/--	2-5,14, 22-31
<input checked="" type="checkbox"/> Voir la suite du cadre C pour la fin de la liste des documents		
<input type="checkbox"/> Les documents de familles de brevets sont indiqués en annexe		
<b>* Catégories spéciales de documents cités:</b> <div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent</p> <p>"E" document antérieur, mais publié à la date de dépôt international ou après cette date</p> <p>"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)</p> <p>"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens</p> <p>"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée</p> </div> <div style="width: 45%;"> <p>"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention</p> <p>"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément</p> <p>"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier</p> <p>"&amp;" document qui fait partie de la même famille de brevets</p> </div> </div>		
Date à laquelle la recherche internationale a été effectivement achevée 13 octobre 2000		Date d'expédition du présent rapport de recherche internationale 23.11.00
Nom et adresse postale de l'administration chargée de la recherche internationale Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Fonctionnaire autorisé CHAMBONNET, F



C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	<p>EMBASE : AC AQ061641;            Créée le 03-AUG-1998 (Rel. 56, Created)            Caractérisation de la séquence :            "CIT-HSP-2348011.TF CIT-HSP Homo sapiens            genomic clone 2348011, Genomic survey            sequence.Homo sapiens (human)"            XP002132840            le document en entier</p> <p>---</p>	1-5
A	<p>RUST, S. ET AL.: "Assignment of Tangier            disease to chromosome 9q31 by a graphical            linkage exclusion strategy."            NAT GENET.,            vol. 20, no. 1, septembre 1998 (1998-09),            pages 96-98, XP000884511            cité dans la demande            le document en entier</p> <p>---</p>	1
A	<p>LUCIANI M F ET AL: "CLONING OF TWO NOVEL            ABC TRANSPORTERS MAPPING ON HUMAN            CHROMOSOME 9"            GENOMICS,US,ACADEMIC PRESS, SAN DIEGO,            vol. 21, no. 1, 1 mai 1994 (1994-05-01),            pages 150-159, XP000869719            ISSN: 0888-7543            cité dans la demande            le document en entier</p> <p>---</p>	1
A	<p>LISCUM L, MUNN NJ.: "Intracellular            cholesterol transport."            BIOCHIM BIOPHYS ACTA.,            vol. 1438, no. 1,            19 avril 1999 (1999-04-19), pages 19-37,            XP000889761            page 33, colonne 1, alinéa 5.2 -colonne 2</p> <p>---</p>	1
P,X	<p>GURA, T.: "Gene linked to faulty            cholesterol transport."            SCIENCE.,            vol. 285, no. 5429,            6 août 1999 (1999-08-06), pages 814-815,            XP000877245            le document en entier</p> <p>---</p>	1
P,X	<p>RUST S, ROSIER M, FUNKE H, REAL J, AMOURA            Z, PIETTE JC, DELEUZE JF, BREWER HB,            DUVERGER N, DENEFLÉ P, ASSMANN G.:            "Tangier disease is caused by mutations in            the gene encoding ATP-binding cassette            transporter 1."            NAT GENET.,            vol. 22, no. 4, août 1999 (1999-08), pages            352-355, XP000884993            le document en entier</p> <p>---</p>	1-5
	<p>---</p> <p>-/--</p>	

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
P,X	REMALEY AT, ET AL.: "Human ATP-binding cassette transporter 1 (ABCI): genomic organization and identification of the genetic defect in the original Tangier disease kindred." PROC NATL ACAD SCI U S A., vol. 96, no. 22, 26 octobre 1999 (1999-10-26), pages 12685-12690, XP000877247 le document en entier ---	1-5
P,X	LAWN RM, WADE DP, GARVIN MR, WANG X, SCHWARTZ K, PORTER JG, SEILHAMER JJ, VAUGHAN AM, ORAM JF.: "The Tangier disease gene product ABC1 controls the cellular apolipoprotein-mediated lipid removal pathway." J CLIN INVEST., vol. 104, no. 8, octobre 1999 (1999-10), pages R25-R31, XP000884782 le document en entier ---	1-5
P,X	BODZIOCH, M. ET AL.: "The gene encoding ATP-binding cassette transporter 1 is mutated in Tangier disease." NAT GENET., vol. 22, no. 4, août 1999 (1999-08), pages 347-351, XP000889766 le document en entier ---	1
P,X	MARCIL M, ET AL.: "Mutations in the ABC1 gene in familial HDL deficiency with defective cholesterol efflux." LANCET, vol. 354, no. 9187, 16 octobre 1999 (1999-10-16), pages 1341-1346, XP000877242 le document en entier ---	1
P,X	ORSO, E. ET AL.: "Transport of lipids from golgi to plasma membrane is defective in Tangier disease patients and Abcl-deficient mice." NAT GENET. 2000 FEB;24(2):192-6., XP000889762 le document en entier ---	1
P,X	BROOKS-WILSON A, ET AL.: "Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency." NAT GENET, vol. 22, no. 4, août 1999 (1999-08), pages 336-345, XP000889767 le document en entier -----	1

**Cadre I Observations - lorsqu'il a été estimé que certaines revendications ne pouvaient pas faire l'objet d'une recherche (suite du point 1 de la première feuille)**

Conformément à l'article 17.2)a), certaines revendications n'ont pas fait l'objet d'une recherche pour les motifs suivants:

1. ☐ Les revendications n<sup>os</sup> se rapportent à un objet à l'égard duquel l'administration n'est pas tenue de procéder à la recherche, à savoir:
2. ☐ Les revendications n<sup>os</sup> se rapportent à des parties de la demande internationale qui ne remplissent pas suffisamment les conditions prescrites pour qu'une recherche significative puisse être effectuée, en particulier:
3. ☐ Les revendications n<sup>os</sup> sont des revendications dépendantes et ne sont pas rédigées conformément aux dispositions de la deuxième et de la troisième phrases de la règle 6.4.a).

**Cadre II Observations - lorsqu'il y a absence d'unité de l'invention (suite du point 2 de la première feuille)**

L'administration chargée de la recherche internationale a trouvé plusieurs inventions dans la demande internationale, à savoir:

voir feuille supplémentaire

1. ☐ Comme toutes les taxes additionnelles ont été payées dans les délais par le déposant, le présent rapport de recherche internationale porte sur toutes les revendications pouvant faire l'objet d'une recherche.
2. ☐ Comme toutes les recherches portant sur les revendications qui s'y prêtaient ont pu être effectuées sans effort particulier justifiant une taxe additionnelle, l'administration n'a sollicité le paiement d'aucune taxe de cette nature.
3. ☐ Comme une partie seulement des taxes additionnelles demandées a été payée dans les délais par le déposant, le présent rapport de recherche internationale ne porte que sur les revendications pour lesquelles les taxes ont été payées, à savoir les revendications n<sup>os</sup>
4. ☒ Aucune taxe additionnelle demandée n'a été payée dans les délais par le déposant. En conséquence, le présent rapport de recherche internationale ne porte que sur l'invention mentionnée en premier lieu dans les revendications; elle est couverte par les revendications n<sup>os</sup> revendications : 1-5, 14-15, 22-34, 41-45 toutes partiellement

Remarque quant à la réserve

- ☐ Les taxes additionnelles étaient accompagnées d'une réserve de la part du déposant.
- ☐ Le paiement des taxes additionnelles n'était assorti d'aucune réserve.

## SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

L'administration chargée de la recherche internationale a trouvé plusieurs (groupes d') inventions dans la demande internationale, à savoir:

1. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 1, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucleotide de séquence SEQ ID NO 15, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un fragment SEQ ID NO 48, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucleotide de séquence nucléotidique SEQ ID No 48, et les acides nucléiques de séquence complémentaire à ces polynucleotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

2. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 2, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucleotide choisi parmi les séquences SEQ ID NO 16-21, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucleotide choisi parmi les séquences SEQ ID NO 49-54, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucleotide choisi parmi les séquences nucléotidiques SEQ ID No 49-54, et les acides nucléiques de séquence complémentaire à ces polynucleotides; sonde ou amorce nucléotidiquespécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

3. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 3, ou un acide nucléique de séquence complémentaire; acide nucléique

## SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

comprenant un polynucleotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 22, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un fragment SEQ ID NO 55, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide de séquence nucléotidique SEQ ID No 55, et les acides nucléiques de séquence complémentaire à ces polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

## 4. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 4, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucléotide de séquence SEQ ID NO 23, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucleotide choisi parmi les séquences SEQ ID NO 56 et 57, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide choisi parmi les séquences SEQ ID NO 56 et 57, et les acides nucléiques de séquence complémentaire à ces polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

## 5. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 5, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucleotide choisi parmi les séquences SEQ ID NO 24-30, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucleotide choisi parmi les séquences SEQ ID NO 58-65, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID No 58-65, et les acides nucléiques de séquence complémentaire à ces

## SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA 210

polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

## 6. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 6, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucleotide choisi parmi les séquences SEQ ID NO 31-34, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucleotide choisi parmi les séquences SEQ ID NO 66-70, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID No 66-70, et les acides nucléiques de séquence complémentaire à ces polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

## 7. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 7, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucleotide choisi parmi les séquences SEQ ID NO 35-38, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucleotide choisi parmi les séquences SEQ ID NO 71-75, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID No 71-75, et les acides nucléiques de séquence complémentaire à ces polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de

## SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

composition pharmaceutique;

8. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 8, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucleotide de séquence SEQ ID NO 39, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucleotide choisi parmi les séquences SEQ ID NO 76 et 77, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID No 76 et 77, et les acides nucléiques de séquence complémentaire à ces polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

9. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 9, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucleotide de séquence SEQ ID NO 40, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucleotide choisi parmi les séquences SEQ ID NO 71-75, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID No 71-75, et les acides nucléiques de séquence complémentaire à ces polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

10. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 10, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucleotide choisi parmi les séquences SEQ

## SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

ID NO 41 à 42, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucleotide choisi parmi les séquences SEQ ID NO 80-82, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID No 81-82, et les acides nucléiques de séquence complémentaire à ces polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

## 11. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 11, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucleotide choisi parmi les séquences SEQ ID NO 43 et 44, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucleotide choisi parmi les séquences SEQ ID NO 83 et 84, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID No 83 et 84, et les acides nucléiques de séquence complémentaire à ces polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

## 12. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 12, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucleotide de séquence SEQ ID NO 45, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucleotide choisi parmi les séquences SEQ ID NO 86 et 87, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID No 86 et 87, et les acides nucléiques



## SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

de séquence complémentaire à ces polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

## 13. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 13, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucleotide de séquence SEQ ID NO 46, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucleotide choisi parmi les séquences SEQ ID NO 88 et 89, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide choisi parmi les séquences nucléotidiques SEQ ID No 88 et 89, et les acides nucléiques de séquence complémentaire à ces polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

## 14. revendications: 1-5, 14-15, 22-34, 41-45, toutes partiellement

Acide nucléique comprenant au moins 245 nucléotides consécutifs d'une séquence nucléotidique SEQ ID No 14, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant un polynucleotide de séquence SEQ ID NO 47, ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins 8 nucléotides consécutifs d'un fragment SEQ ID NO 90, ou un acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide de séquence nucléotidique SEQ ID No 90, et les acides nucléiques de séquence complémentaire à ces polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique;

## SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

## 15. revendications: 6-9 et partiellement 14-15, 22-34, 41-45 50-51

Acide nucléique comprenant polynucléotide de séquence SEQ ID NO 91 ou un acide nucléique de séquence complémentaire; acide nucléique comprenant au moins huit nucléotides consécutifs d'un polynucléotide de séquence nucléotidique SEQ ID No 92, un fragment biologiquement actif de celui-ci ou un acide nucléique de séquence complémentaire, ou un homologue de celui-ci; sonde ou amorce nucléotidique et les acides nucléiques de séquence complémentaire à ces polynucléotides; sonde ou amorce nucléotidique spécifique du gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi de tels acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; utilisation desdits acide nucléique ou vecteur pour la fabrication de composition pharmaceutique; utilisation de ladite cellule pour cribler des principes actifs pour la prévention ou le traitement de maladies résultant d'un dysfonctionnement du transport inverse du cholestérol;

## 16. revendications: 35 et partiellement 10-12, 16-17, 22 -34, 37-40

Acide nucléique ayant au moins huit nucléotides consécutifs d'un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 93-94 ou un acide nucléique de séquence complémentaire; acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 93-94, ou un acide nucléique de séquence complémentaire; sonde ou amorce nucléotidique pour la détection d'une mutation dans le gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi ces acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; polypeptide ABC1 muté, caractérisé en ce qu'il comprend un polypeptide de séquence d'acides aminés SEQ ID No 140; anticorps dirigé contre ce polypeptide ainsi que les procédé de détection l'utilisant et nécessaire le comprenant;

## 17. revendications: 36 et partiellement 10-12, 16-17, 22 -34, 37-40

Acide nucléique ayant au moins huit nucléotides consécutifs d'un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 95-96 ou un acide nucléique de séquence complémentaire; acide nucléique ayant au moins 80% d'identité en nucléotides ou hybridant dans des conditions de forte stringence avec un polynucléotide choisi dans le groupe constitué des séquences nucléotidiques SEQ ID NO 95-98, ou un acide nucléique de séquence complémentaire;

## SUITE DES RENSEIGNEMENTS INDIQUES SUR PCT/ISA/ 210

sonde ou amorce nucléotidique pour la détection d'une mutation dans le gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi ces acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique; polypeptide ABC1 muté, caractérisé en ce qu'il comprend un polypeptide de séquence d'acides aminés SEQ ID No 141; anticorps dirigé contre ce polypeptide ainsi que les procédé de détection l'utilisant et nécessaire le comprenant;

## 18. revendications: 10-15, 18-34, toutes partiellement

Acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucléotide de séquence SEQ ID NO 97 ou 98 et comprenant la base polymorphe ou un acide nucléique de séquence complémentaire; sonde ou amorce nucléotidique pour la détection d'un polymorphisme, dans le gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi ces acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique;

## 19. revendications: 10-15, 18-34, toutes partiellement

Acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucléotide de séquence SEQ ID NO 99 ou 100 et comprenant la base polymorphe ou un acide nucléique de séquence complémentaire; sonde ou amorce nucléotidique pour la détection d'un polymorphisme dans le gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi ces acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique;

## 20. revendications: 10-15, 18-34, toutes partiellement

Acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucléotide de séquence SEQ ID NO 101 ou 102 et comprenant la base polymorphe ou un acide nucléique de séquence complémentaire; sonde ou amorce nucléotidique pour la détection d'un polymorphisme dans le gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi ces acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique;

## 21. revendications: 10-15, 18-34, toutes partiellement

## SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

Acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucléotide de séquence SEQ ID NO 1037 ou 104 et comprenant la base polymorphe ou un acide nucléique de séquence complémentaire; sonde ou amorce nucléotidique pour la détection d'un polymorphisme dans le gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi ces acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique;

## 22. revendications: 10-15, 18-34, toutes partiellement

Acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucléotide de séquence SEQ ID NO 105 ou 106 et comprenant la base polymorphe ou un acide nucléique de séquence complémentaire; sonde ou amorce nucléotidique pour la détection d'un polymorphisme dans le gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi ces acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique;

## 23. revendications: 10-15, 18-34, toutes partiellement

Acide nucléique comprenant au moins 8 nucléotides consécutifs d'un polynucléotide de séquence SEQ ID NO 107 ou 108 et comprenant la base polymorphe ou un acide nucléique de séquence complémentaire; sonde ou amorce nucléotidique pour la détection d'un polymorphisme dans le gène ABC1 d'une longueur d'au moins 15 nucléotides choisie parmi ces acides nucléiques ainsi que les procédés d'amplification et de détection les utilisant et leurs nécessaires de mise en oeuvre; vecteur et cellule hôte recombinants comprenant un tel acide nucléique;

## 24. revendications: 46,47

Utilisation du polypeptide ABC1 ou de cellules exprimant ce polypeptide pour la fabrication d'un médicament destiné à la prévention de l'athérosclérose et composition pharmaceutique dérivée;

## 25. revendications: 48-49 et partiellement 50-51

Utilisation du polypeptide ABC1 ou de cellules exprimant le polypeptide ABC1 pour cribler des principes actifs pour la prévention ou le traitement de maladies résultant d'un dysfonctionnement du transport inverse du cholestérol, pour

SUITE DES RENSEIGNEMENTS INDICES SUR PCT/ISA/ 210

autant que non couvert par le sujet 15.